Possibilistic c-Means -Spatial Contexutal Information based sub-pixel classification approach for multi-spectral data

Sivee Chawla

January, 2010

Possibilistic c-Means -Spatial Contexutal Information based sub-pixel classification approach for multi-spectral data

by

Sivee Chawla

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree in Master of Science in *Geoinformatics*.

Degree Assessment Board

Thesis advisor	Dr. Anil Kumar Shri. PLN Raju Dr. Nicholas Hamm Dr. V.A Tolpekin
Thesis examiners	Prof. Dr. V.G. Jetten (Chairman) Dr. Nicholas Hamm (ITC,Member) Prof. Dr. P.K Garg (External Examinar) Sh. P L N Raju (IIRS, Member) Dr. Sameer Saren (IIRS, Member)

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS
AND
INDIAN INSTITUTE OF REMOTE SENSING
DEHRADUN, INDIA

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation (ITC). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

Since, the advent of remote sensing technology it has become an attractive option to capture the information about land cover classes at a global scale in less time in form of digital images. These images are available over a large range of temporal and spatial scales. But the complete utilization of this information is dependent on the availability of efficient and accurate methods of image analysis. Image classification is one of the important image analysis methods.

The efficiency of image classification to map the real world scenario is restricted by the presence of mixed pixels. To some extent fuzzy based classification methods such as Fuzzy-c-Means or FCM, can handle these mixed pixels. FCM gives the membership value of the pixel to various classes. But the accuracy of this approach is found to be less than that of Possibilistic c-Means (PCM). The membership value generated by PCM is interpreted as the degree of belongingness instead of degree of sharing i.e. generated by FCM.

But PCM exploits the information present only in spectral domain. Where as, it has been observed that including the information from spatial domain increases the accuracy of the classifier. This spatial domain information or spatial contextual information can be integrated with PCM to provide more accurate results.

In this research work, a PCM and spatial contextual information based sub-pixel classification method has been developed. Also, two different MRF prior models have been used. It has been observed that using Discontinuity Adaptive prior models preserve the edges and also, increase the overall accuracy of the classifier.

It has been found with the experiments performed, that not all sub-pixel accuracy assessment techniques are suitable in case of PCM classifier such as PCM-MRF, as it doesn't follow the hyperline constraint. For generating the reference data set, a finer resolution image was used. It has also been discussed with results, that same classifier is more appropriate to use for generating the reference data set.

Keywords

Possibilistic -c-Means, Markov Random Fields, DA models, edge preservation

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Many countries all around the world, including India, now depend on remotely sensed data and images for monitoring, planning and development of natural and man-made resources.

For the preparation of thematic maps and quantitative analysis of the images, capability of computer to interpret the images, identify pixels and label them based on their numerical properties is exploited [1]. The method is commonly referred to as *Image Classification*. And as the images are acquired and process in digital format it is better known as Digital Image Classification.

Conventional classification methods assume that the area under study is composed of number of unique, internally homogenous classes that are mutually exclusive [2]. Also known as hard classification, it follows the classical set theory i.e. if land cover classes are assumed to be set of pixels, then a pixel will belong to only one such set and have zero membership to all other sets. Mapping the same in feature space, where feature space is partitioned into clusters, which are made up of points representing DN values of individual pixels in the image. As per hard classification, the pixel will belong to either one of these clusters (also, known as spectral classes /information classes, accordingly), thus, partitions will have defined boundaries in the feature space. But this is not always the case. The pixels may belong to two or more classes. In real world, the pixels are large enough to accommodate sub-pixel objects such as trees, sub-pixel linear objects viz. rivers, boundaries of more than two land cover types or intergrade phenomena may be observed (figure 1.1) [3]. That is the pixels are mixed in nature and not pure, in terms of classes they cover. This commands a need of classifiers which can handle mixed pixels by giving membership value to the pixels in different classes. Thus, providing classified outputs which are near to real life scenarios.

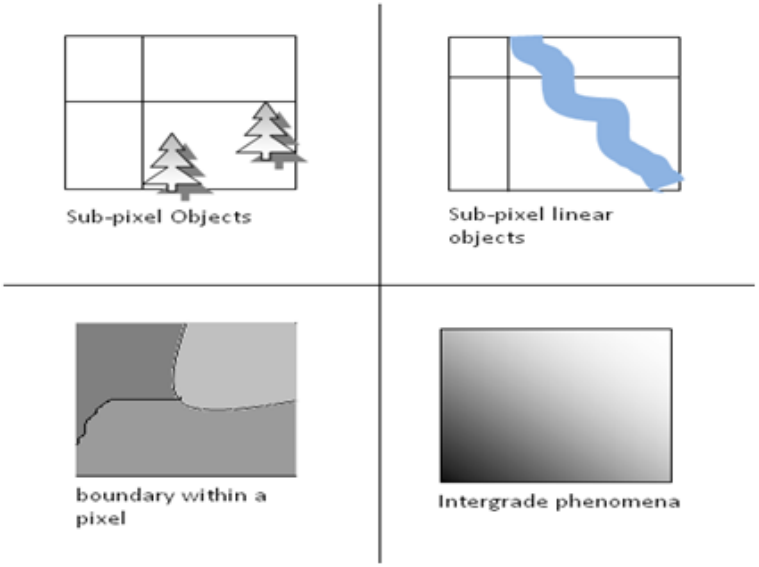


Figure 1.1: The sub-pixel issues Based on (3)

1.2 The fuzzy classifier

1.2.1 The Problem

As described by Zhang and Foody [2], the reason behind mixed pixels can be mainly attributed (there can be other reasons, though) to following three sources:-

1. In real world, different land cover types are rarely homogenous but are heterogeneous resulting in intergrade phenomena i.e. classes do not have crisp boundaries. This is fuzziness due to the geographical phenomena.

2. A pixel value is the result of interaction of electromagnetic waves with the ground objects and/or atmosphere. The sensor records this spectral response which may differ for similar entities, while dissimilar entities may show similar spectral response, depending on the ground situations.

3. Finally, due to coarse or medium spatial resolution of sensors (e.g. AWIFS and LISS III), a pixel may not consist of single class but two or more classes irrespective of the fact that on the ground the class may be homogenous. This can be defined as fuzziness due to resolution of the sensor.

1.2.2 The Solution(s)

Fuzzy set theory [4] is one of the possible ways which takes care of the above problems to an extent. It assigns membership value to a pixel, in different classes. These membership values can take any real number between zero and one. Fuzzy c-Means (FCM) [7] is a popular fuzzy set theory based sot classifier, which handles the vagueness of a pixel at sub-pixel level.

FCM has been successful in assigning the membership (u_{ij}) of a pixel to multiple classes but the assignment is relative to total number of classes defined [6, 5]. This is because of the constraint on the membership values given by the equation 1.1:-

$$\sum_{i=1}^{C} u_{ij} = 1, \forall j \tag{1.1}$$

where, j varies from 1 to N (N is the total no. of pixels in the image) and C is the total no. of classes defined by the analyst.

Equation 1.1 can be interpreted as, sum of membership values for a pixel in all the classes should be equal to one [6, 7].

Krisnhapuram and Keller in 1993, gave a variation of FCM, called Possibilistic *c*-Means (PCM) which relaxes the constraint on membership value in equation 1.1, ensuring the condition given in equation 1.2 [6] :

$$\max_{i} \; u_{ij} > 0, \quad \forall j \tag{1.2}$$

In case of PCM, this membership value represents the "degree of belongingness or compatibility or typicality"[6], contrary to that represented by FCM , where it is , "degree of sharing"[6]. Overcoming this constraint gives higher accuracy of supervised classification using PCM as compared to that of FCM[30, 29]. Also, PCM can handle noise and outliers [6, 11]. Noise and outliers affect the prototype parameters i.e. cluster means. Also, PCM, as a supervised classifier, works better in case of untrained classes , when compared with FCM (as supervised classifier)[5]. Untrained classes here are those classes which are present in the image but are not known to the analyst; hence, the classifier is not trained with that unknown class.

Thus, the advantages of PCM over FCM are the motivation behind selecting PCM as sub-pixel classification approach in this research.

1.3 Contextual Information and Markov Random Fields

In traditional classification methods, the classiification is performed based only on the pixel's spectral value. These methods assume that the spectral properties of a pixel are independent of the properties of pixels in the neighbourhood. This situation seldom exists in the actual geographical phenomena. On the contrary, there is a correlation between a pixel and its adjacent pixels in the image space. This happens primarliy due to two reasons [1]:

1. The reflected energy reaching the sensor is not solely from the pixel but also, consists of energy from neighbourhood pixels as well.

2. The classes on the ground are usually, large enough to cover more than one pixels and in fact, they cover at least few pixels. Thus, hardly, any pixel exists in isolation. A pixel belonging to a forest has a high probability to have forest pixels as its neighbours.

The information from neighbouring pixels (in image space) is known as spatial contextual information.

Spatial contextual information, if exploited properly, may allow the elimination of ambiguities (such as two pixels may have similar spectral values but belong to different classes that can be recognized only with the context), recovery of missing information and the correction of errors [8], and thus, can provide spatially and spectrally consistent thematic maps.

The methods to add context are categorized into three broad categories [1]:

1. **Image pre-processing** In this method, the contextual information is dealt-with before classification, by modifying or enhancing the spatial properties of the pixel.Various methods include *median filters, averaging filters, ECHO classification methodology, generating seperating channels such as texture channel.*

2. **Post-Classification** After the classification, a filter such as *Majority Filter, Thomas Filter* [14] can be applied over the classified image.

3. **Label Relaxation** This method incorporates spatial contextual information into classification process in a logically consistent way using a methodological framework [17]. Markov Random Fields (MRF) is based on the Hammersley Clifford Theorem of MRF-GRF equivalence. It is an attractive option as it provides a flexible stochastic framework for modelling a given scene, expressing spatial contextual information by means of adequate energy function[9] . Since, classical paper by Geman & Geman in 1984 [10], MRF (with Simulated Annealing) has become a popular and powerful tool for incorporating spatial contextual information in the images for processes such as segmentation, restoration of noisy images, classification etc.

1.3.1 Edge Preserving priors

Generally, MRF assumes smoothness or homogeneity in the images but its improper use can lead to undesirable over-smoothing. This usually occurs at the discontinuities such as edges. Therefore, it is important to select MRF models (also called MRF priors and regularizers) which take into account the discontinuities. The MRF models include various regularizers such as *standar regularization model, line process, weak string & memberane and Discontinuity Adaptive (DA) smoothness models which focus on first order derivatives* [21].

In this thesis, two MRF models have been used, one is standard regularization model and other is Discontinuity Adaptive smoothness priors.

1.4 Problem Statement

Fuzzy based techniques such as FCM are used for sub-pixel classification of images, giving higher accuracy over the hard classifiers as it incorporates vagueness of classes. An even higher accuracy has been registered by using PCM as compared to FCM in sub-pixel classification methods [28, 29, 30]. It has also been shown that in case of hard classification algorithm [8] as well as in case of FCM [12, 13], that if the spatial contextual information is included then it becomes more robust as compared to purely spectral based classification [1].

As PCM work only in spectral domain thus, a PCM and spatial contextual information based sub-pixel classification method has been developed, that takes into account the fuzziness of classes and also, incorporates spatial contextual information.

1.5 Research Identification

To achieve the solution of the problem statement defined in sectin 1.4, following research objectives and corresponding research questions, have been formulated (Table 1.1).

Table 1.1: Research Objectives and Questions

Research Objectives	**Research Questions**
I)To develop a PCM based sub-pixel classifier incorporating spatial contextual information using MRF that preserve edges.	1) How to modify the PCM objective function so as to include the spatial contextual information using MRF? 2)Which smoothness prior model will be suitable for PCM-MRF ? 3) Which sampler will be suitable for PCM-MRF?
II) To assess the performance of PCM-MRF in presence of untrained classes noise and outliers.	4) Which classifier to be used for classifying finer resolution images i.e. LISS IV images, to generate the reference data set? 5) Which accuracy assesment tech--nique is suitable to assess the accuracy of PCM-MRF sub-pixel classifier? 6) Which accuracy assessment technique will be suitable to assess the accuracy of PCM-MRF at edges? 7) How does PCM-MRF performs in case of noise/outliers/ untrained classes?

1.6 Research Setup

To answer the research questions and fulfill research objectives, the layout of the steps followed in this research thesis is given in Figure 1.2.

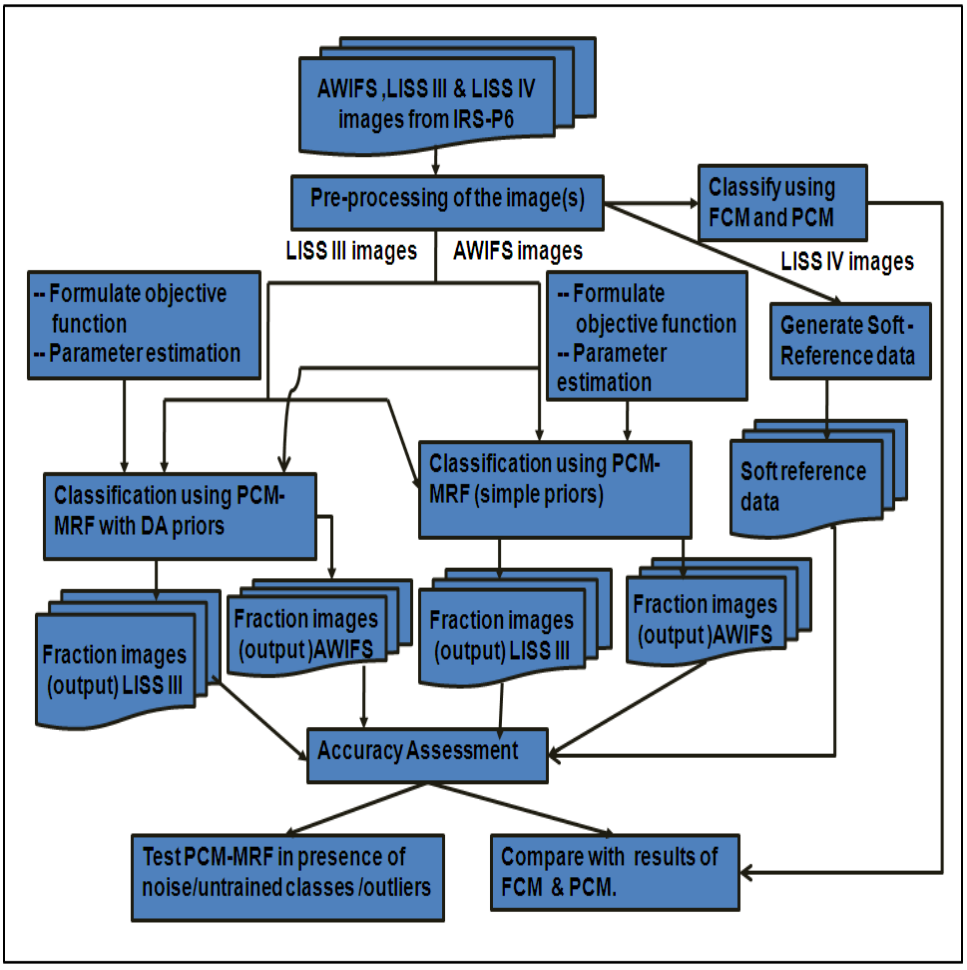


Figure 1.2: Flow Diagram of the complete research setup

1.7 Thesis Structure

The whole thesis has been organised into eight chapters. The ***first chapter*** includes *background* information on the major aspects of the research topic, the *problem statement*, *research objectives and questions* and *research setup*. In ***Second Chapter*** literature review of the research done in the topics relevant to this thesis is givens. ***Third and Fourth Chapter*** throws light on theory behind PCM and MRF, respectively. ***Fifth chapter*** provides information on Study Area and the data set used. ***Sixth chapter*** has been dedicated to details of the complete methodology. Followed by results and their discussion in ***seventh chapter***. Finally, in ***eighth chapter*** the thesis has been concluded with recommendations pertaining to further research in the area.

Chapter 2

Literature Review

2.1 Possibilistic c-Means and fuzzy based clustering methods

Krishnapuram & Keller [6] gave a specific implementation of Zadeh's [4] possibility based method, called Possibilistic c-means. It assigns pixel to more than one cluster in the form of membership value and this membership value doesn't follow the constraint in FCM (equation 1.1) called hyper-line constraint [6].

Krishnapuram & Keller [11] gave some recommendations based on their findings and issues raised by Barni et al. In this paper, they said that 'fuzzifier (m)' is different for both FCM and PCM as its interpretation is differs. Another parameter in PCM i.e η_i was defined as the distance at which membership value becomes 0.5 .

Foody [5] showed that supervised PCM in case of untrained classes (explained in section 3.2.1, chapter3) gave lower RMS error and r, when compared to that of Fuzzy C-Means.

Kumar et al. [31] have given the comparison of FCM vs. PCM as sub-pixel methods and results show that PCM with Euclidean Norm had the highest overall accuracy of 99.29%, whereas FCM with Euclidean Norm showed the overall accuracy of 97.9%.

In another work by Kumar et al[29], in which PCM and FCM were applied first on first three bands and then later on first nine bands, followed by 14 bands of ASTER data. The results on all three cases gave the higher performance of PCM as compared to FCM in terms of accuracy. While overall accuracy of FCM in case of 3 bands was 76.4%, it was 86% for PCM. In case of 9 bands also, when FCM was 81.7%, the overall accuracy of PCM was 88.9%.

The performance of PCM has been shown to be better than FCM [5, 31, 29] because PCM overcomes the hyperline constraint in FCM.

2.2 Markov Random Field and Image Analysis

Markov Random Field has been used in image analysis for both medical and Remote Sensing images. It has been used for including contextual information, particularly, spatial contextual information.

The seminal paper by Geman and Geman [10], showed the implementation of Simulated Annealing using Gibbs Sampler for MRF in image restoration. Following this MRF has been used extensively in image analysis.

Solberg et al[17] proposed a MRF based model for classification of remotely sensed data obtained from multiple sources. They exploited both spatial and temporal contextual information, successfully. When multi-temporal MRF fusion model was used, the overall accuracy imrovement was reported to be 2.7% as compared to reference model. And with the inclusion of crop field border map from GIS data, there was a significant increase in the overall accuracy, which was 79.6%.

Pham [12] included spatial contextual information with FCM using MRF for image segmentation in Magenetic Resonance Images (MRI) of brain and called it as Robust Fuzzy C-means or RFCM Algorithm. Convergence of the objective function was achieved when change in the objective function was less than a defined threshold. To obtain the value of β which controls the smoothness performed by the penalty function (or objective function), cross-validation technique used. The results were compared using misclassified rate (MCR) which was 14.14% for FCM, whereas for RFCM it was 0.52%.

Binaghi et al [18] integrated the Fuzzy based knowledge classifier with contextual information using MRF and remote sensing data from multiple sources. The attempt was emulate the process of human thinking in solving problems pertaining to classification.

2.2.1 Preventing Oversmoothning

MRF use smoothness priors to include spatial contextual information and to avoid over smoothning, Regularizers and Discontinuity Adaptive (DA) Models have been introduced.

Li [22, 23], gave DA models to be used as prior models in MRF, which are said to take into account the discontinuities and avoid over smoothning. In [23], it was shown that solution to DA models can be obtained by using gradient descent method, but its direct use may cause getting trapped into local minima. Further details are also, provided in [21].

Smits and Dellepiane [34] gave a 'Discontinuity Adaptive MRF' model for segmentation of SAR images using unsupervised mode of segmentation. Because it is important to preserve the discontinuities (in particular, small structures) in SAR images as it is obscured by speckle noise [34].

Tso & Olsen [32] developed a method for classifying multispectral images by incorporating contextual information and also including edge information. The edge information was extracted from high resolution panchromatic image at multiple scales and then later on fused using fuzzy line process models with contextual information. The authors used ICM for energy minimization. They also stated that at boundary the MRF tends to over-smooth the image and causes introduction of error. Thus, to avoid errors it is important to identify and preserve the edges and reduce their contribution in classification, particularly in smoothness prior model.

Hou et al[35], used modified FCM that includes spatial contextual infor-

mation using moving average filter based on [12] as the regularizer. An accuracy was tested using different noise levels and similar results as seen in [12], were obtained with reduced time complexity by 30%. Thus, they gave another method of including regularizer in FCM based clustering for segmentation.

As seen in the literature survey, MRF has been included into fuzzy based approaches but use of MRF with PCM for sub-pixel classification has not been in seen in the literature, especially in the field of remote sensing image. Neither edge preserving models for MRF with PCM has been used for soft classification.

2.3 Validation Techniques

Accuracy assessment and validation for sub-pixel classifiers is still a subject of research. No standard methods are available for sub-pixel classifiers, unlike that in hard-classifiers such as Error Matrix and Kappa Coefficient.

In 2006, Pontius and Cheuk[46] gave a set of composite operators for computation of cross-tabulation matrix for soft classified outputs. These composite operators are combination of single operators including boolean, multiplication and minimum operators. These operators are used in building cross-tabulation matrix and assess the results for the comparison of classifier.

Another measure that can be used to provide the quantitative measure of the reliability of classification is entropy. Ricotta and Avena [48] described a generalized function of entropy that is sensitivity to the presence of abundant class and a rare one.

Dehghan and Ghassemian [47] have given the entropy as an absolute measure of uncertainty for the classified output. It is called absolute because it doesn't take into account any other reference data set like in case of RMSE and correlation coefficient. Also, it states that a single number can be used to give the uncertainty of the classified output at per pixel level or per class level or even at image level. It states, that higher entropy implies higher uncertainty and vice-versa. A classifier with lower entropy is a better classifier.

Chapter 3

The Possibilstic *c*-Means

"Building Insights.Breaking Boundaries." -Elseiver

This chapter is about what and why of PCM.'What' implies, what is actually PCM and 'Why' here tells, why it has been selected over FCM.

Based on the possibility theory given by [4], Krishnapuram and Keller [6] gave a particular implementation of the possibility theory, termed as Possibilistic *c*-Means(PCM). It is a clustering algorithm, which is the modification over FCM objective function.

In the following sections, only those terms pertinent to the current thesis work have been covered. Section 3.1 talks about FCM, which in turn forms the basis of discussion for the central theme of the chapter i.e. PCM in section 3.2. Thus, in section 3.1 and 3.2 an 'insight is built' into what and how of PCM and the mathematical constraint overcome by PCM is all about 'breaking boundaries '.

3.1 Fuzzy C-Means

FCM, is primarily a partitioning algorithm [11]. In terms of image processing, it'll partition the feature space into number of clusters defined by the user (figure 3.1). This is regardless of actual number of clusters present in the data set. It is mainly due to the probabilistic constraint on the membership value mentioned in equation 1.1, chapter 1.

The membership value generated using FCM gives the *"degree of sharing"* of the pixel in the various clusters in the feature space. In other words, is the degree to which a pixel will be shared among the clusters. It can be better explained by taking up a real life example.

Example 3.1 *Assuming that there is a tub of pudding that has to be shared among a group of people. The constraint here is that stock of pudding is limited and will be shared among given set of people. The size of the share will depend upon total number of people in the group present, other factors such as ones liking and disliking are kept in the background for the time being*

The share here can be compared to the *membership value of the pixel* and the constraint, in case of FCM, is *the sum of membership value of the pixel should be equal to one* (equation 1.1, chapter 1).

It has been observed that due to this constraint FCM does nots perform well in case of noise and outliers as it affects the formation of clusters [6, 11]. Also, in case of untrained classes (during supervised classification), as demonstrated by [5], FCM

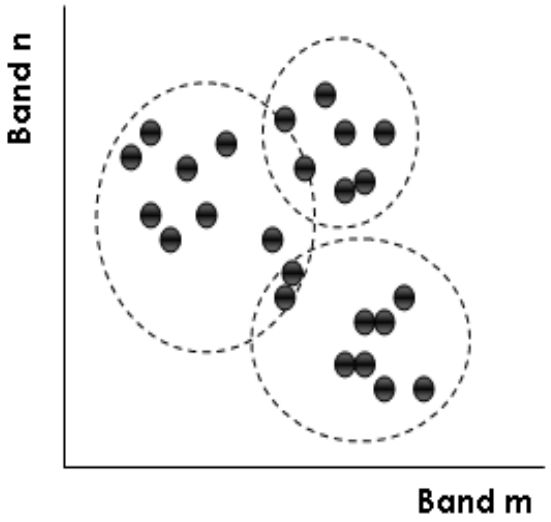


Figure 3.1: Partitioning of feature space using FCM

is less efficient than PCM. It is because it gives the relative membership value and not absolute i.e. the share of pudding one gets depends on the total number of people present.

Hence comes into the picture, PCM.

3.2 Possibilistic *c*-Means

3.2.1 Why PCM?

Unlike FCM, membership value generated by PCM can be interpreted as "*degree of belongingness or compatibility or typicality*"[6]. *Degree of Belongingness* implies the degree to which the pixel belongs to a class. *Degree of compatibility* is the degree by which the pixel is compatible to the other pixels in the cluster and the cluster mean. *Degree of typicality* helps to differentiate between a highly atypical member of the cluster versus moderately atypical member of the cluster. All, the three above give the possibility of a point or pixel to belong to a class. This is contrary to that of FCM, where it is the "degree of sharing"

The above stated difference in the interpretation of membership values in case of FCM and PCM is observed as PCM relaxes the constraint on FCM given in equation 1.1, by following equation 1.2 (given in chapter 1). And this also, improves the performance of PCM in case of:

1. **Outliers**. Outliers are those rare pixels which belong to a class but are at larger distance from the cluster (in feature space) as compared to other pixels of the same class. In other words, they depict an exceptional behaviour that does not comply with the general characteristics of the cluster members.

2. **Noise**. The presence of noise affect the results of FCM and other classification techniques as it influences the estimation of cluster centres [6, 15]. Noise can be defined as anomaly in the set of pixels i.e. those pixels which do not belong to any meaningful cluster.

3. **Untrained Classes**.[5] demonstrated that PCM performs better as compared to FCM in case of untrained classes, during supervised classification. Untrained classes can be defined as those classes which an analyst fails to specify or recognize during the training stage, though they are present in the image.

Referring back to the example of 'pudding', if the tub is bottomless i.e. the quantity of pudding is very large, and then everyone will get enough pudding to satiate one's hunger. Comparing this with the constraint relaxed by PCM (equation 1.2, chapter 1), i.e. if the membership value of a pixel in all the clusters is maximum, then it will give the degree of belongingness of a pixel to different clusters. The share, here, will not depend on total number of people which in the case of PCM are comparable to clusters.

3.2.2 What is PCM? Mathematics of PCM

In the last section, the difference between PCM and FCM was established. There it was shown that PCM works by relaxing the constraint on FCM. But if this relaxation of constraint is applied on the original objective function of FCM, (equation 3.1), it results in a trivial solution[6].

$$J(L,U) = \sum_{i=1}^{C}\sum_{j=1}^{N}(u_{ij})^m d_{ij}^2 \tag{3.1}$$

where,

$L = (\beta_1, \beta_2, ..., \beta_C)$, *is the set of prototypes(cluster centers) for each class c.*

U, *is a* $C \times N$ *matrix known as fuzzy C-partition matrix. Fuzzy C-partion matrix consists of membership values of each pixel in each class, where a row represents one class and consists membership value of the pixels in the class i.*

N, *total number of pixels in the image*

j, *represents a pixel and varies from 1 to N*

d_{ij}^2, *is the distance of feature point* x_j *i.e. the pixel value to prototype* β_i.

And , u_{ij} is calculated as (which is iteratively updated to minimize the objective function equation 3.1):

$$u_{ij} = [\sum_{k=1}^{c}(\frac{d_{ij}^2}{d_{ik}^2})^{\frac{1}{m-1}}]^{-1} \tag{3.2}$$

Thus, Krishnapuram & Keller in 1993 specified the modified objective fucntion for PCM, which is as follows:

$$J_m(L,U) = \sum_{i=1}^{C}\sum_{j=1}^{N}(u_{ij})^m d_{ij}^2 + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-u_{ij})^m. \tag{3.3}$$

Where,

η_i, is called the "bandwidth or scale or resolution" parameter and is estimated from the data. It is the parameter that determines the distance at which the membership value becomes 0.5 [11]. It is calculated as:

$$\eta_i = \frac{\sum_{j=1}^{N} u_{ij}^m d_{ij}^2}{\sum_{j=1}^{N} u_{ij}^m} \tag{3.4}$$

The membership value, in case of PCM, will be calculated from the following equation:

$$u_{ij} = [\sum_{k=1}^{C} \frac{d_{ij}^2}{\eta_i}^{\frac{1}{m-1}}]^{-1} \tag{3.5}$$

The first term in equation 3.3, forces the distances between the feature vectors i.e. the pixels and their prototypes to be as low as possible. Whereas, the second term, demands u_{ij} to be as large as possible, thus, preventing the onset of trivial solution.

The objective function in equation 3.3 and the membership value in equation 3.5, satisfies the criteria of maximization given in equation 1.2 and also, the following fuzzy criterion (also, applicable for FCM) given below.

$$u_{ij} \in [0, 1], \forall i, j. \tag{3.6}$$

$$0 < \sum_{j=1}^{N} u_{ij} \leq 1, \forall i. \tag{3.7}$$

The objective function of PCM (equation 3.3) explained here, will be used in later chapters to explain the formulation of the final objective function for PCM-MRF and thus, is of prime importance in the complete thesis.

Chapter 4

Theory behind Markov Random Field

"He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast." -Leonardo Da Vinci

Contextual information or related information about the pixel other than just the DN value of the pixel itself, constitutes an important aspect of Digital Image Interpretation. Especially, in the case of image labelling problem, a pixel if considered in isolation may lead to missing or incomplete information. It has been seen that including the contextual information for Image Interpretation, particularly for image classification, has proved to be beneficial [24, 17]. It gives additional information which helps in improving the classification accuracy.

Contextual information or context, can be acquired from i) Spectral ii) Temporal iii) Spatial domain. In this thesis, the context from spatial domain has been exploited for the purpose of classification. It is based on the idea that the brightness value of a pixel have statistical dependence on the intentsities of the surrounding pixels, unless the image is nothing but "random noise "[16].

As stated in chapter 1 (section 1.3), MRF provides a logically consistent and mathematically tractable way of establishing the contextual relationship among pixels of the image. MRF for deriving spatial contextual information, has been used widely in classification [17], image denoising [16], image segmentation[12] etc.

The purpose to include this chapter in the thesis is to provide a required theoretical background on MRF. In section 4.1, the basic components that contribute to shape up MAP-MRF framework has been discussed. A complete section, section 4.2, has been dedicated to the Discontinuity Adaptive priors. And the last section of the chapter i.e. section 4.3 is about the MAP-MRF framework. Section 4.4, gives the summary of the complete chapter.

4.1 Components of Markov Random Field

4.1.1 What is Markov Random Field?

Markov Random Field are nothing but undirected graph ,$G = (V, E)$(figure 4.1), where, the node(V) of graph are the sites on which is defined a set of random variables say $X = x_1, x_2, ..., x_m$. Here, m is the number of sites and family X is the random field. These sites or nodes take up a label from the set $\mathcal{L}$. Set $\mathcal{L}$ is dependent on type of

application, e.g., for example for image restoration $\mathcal{L}$={noisy data,significant data}, for classification $\mathcal{L}$ ={agricultural land, water, forest} and so and so forth. And E i.e. edges of the graph show the probabilistic dependency.

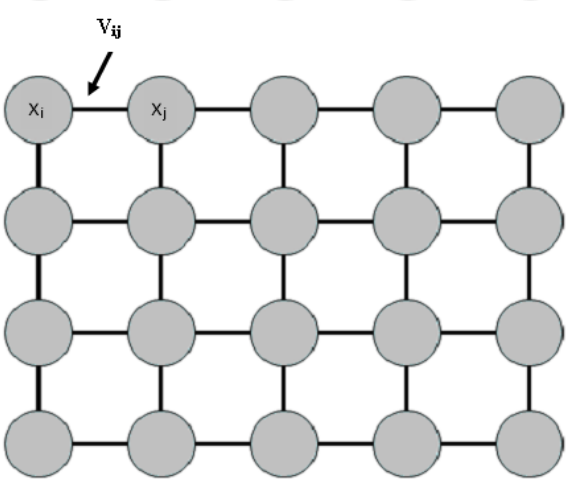


Figure 4.1: Diagram depicting MRF model as an undirected graph

Example 4.1 *Taking up an example from electrical science, the nodes represent the junctions and their label x_j, where $1 \leq j \leq m$, represent the voltage at that point. Visiting figure 4.1 again, and visualizing it as a circuit with junctions. The MRF model here will give how the voltage varies and co-varies among the nodes in the graph or the junction in the circuit. The depedency among the nodes can be captured using the potential term V_{ij}*

***Note**: The term"potential "and other terms such as "temperature"used in MRF, directly come from statistical physics from where MRF intially originated.*

MRF are suitable for establishing a model for probabilistic dependency among the cites or variable,where there is no "apparent directionality to the dependence"[19]. And the potential term V_{ij}, given in figure 4.1 (as well the previous example) is used to establish the probabilistic dependency among nodes, as explained later in the chapter(section 4.1.3 and further).

A formal definition of MRF will be [8] :

Definition 4.1 *For a random field to be a MRF with respect to its neighbourhood, the probability function of MRF should satisfy following properties:-*

1. *Positivity :$P(x) > 0$ for all configurations of x.*
2. *Markovianity : $P(x_j|x_{S-j}) = P(x_j|x_{\mathcal{N}_j})$*
3. *Homogeneity : $P(x_j|x_{\mathcal{N}_j})isthesameforallsitesr$*

where, $S - j$, is the set difference (i.e. all pixels in the set S excluding j). x_{S-j}, denotes the set of labels at the sites in $S - r$. $\mathcal{N}_j$, is the neighbours of the site r.

4.1.2 The window of MRF : Neighbourhood System

Any random field, the probability function of which, satisfies the properties stated in definition 4.1(section 4.1.1) with respect to its neighbourhood is called MRF. The "Neighbourhood System" so much in discussion, can actually be defined as :

Definition 4.2 *$\mathcal{N}_j$ associated with site 'j' is defined as a neighbourhood system if it satisifies following properties :-*

1. $j \notin \mathcal{N}_j$
2. $k \in \mathcal{N}_j \iff j \epsilon \mathcal{N}_k$

Where, k, is from the complete set S of the sites other than target site j.

- **Property 1**, states that the target site i.e. in case of images, the target pixel should not be a part of the neighbourhood system.
- **Property 2**, each should be neighbour of other.

See figure 4.2 for different orders of neighbourhood system.

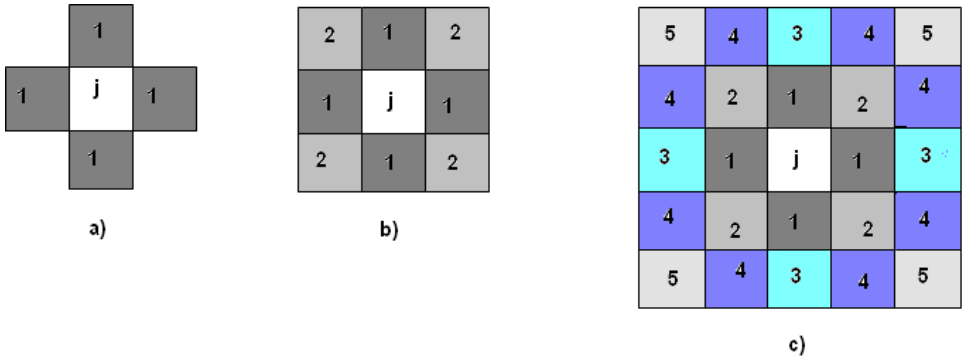


Figure 4.2: Neighbourhood system of different order for pixel j. a) First Order b) Second Order c) Fifth order

For image ,which is a rectangular lattice, first order neighbourhood window will be denoted by $\mathcal{N}(i,j) = [(i-1,j),(i+1,j),(i,j-1),(i,j+1)]$.

This sytem of *neighbourhood window*, allows that the value of nodes in case of MRF to depend only upon the local interactions. Unlike, the global dependency i.e. dependency of every node on the complete set of nodes, in case of Gibbs Random Field (GRF).

4.1.3 GRF-MRF Equivalence

In MRF it is required to establish a probability density function,which is done with the help of "MRF-GRF Equivalence". This MRF-GRF equivalence is given by *Hammersley Clifford Theorem* (for proof refer book by Li[21]), which states that:-

Theorem 4.1 *A random field is a Markov Field for the neighbourhood system X, if and only if, it is a neighbour Gibbs Field for X. [16].*

Also, it states that there exists a unique GRF for every MRF but only if GRF can be defined in terms of clique on a neighbourhood system (see figure 4.2)[8].

As explained in section 4.1.2, GRF considers the global dependency of the nodes. In other words, it gives joint probabilities of the variables in the global context, where Gibbs distribution is given in equation 4.1.

$$P(x) = Z^{-1} \times e^{-\frac{1}{T}U(x)} \tag{4.1}$$

where,

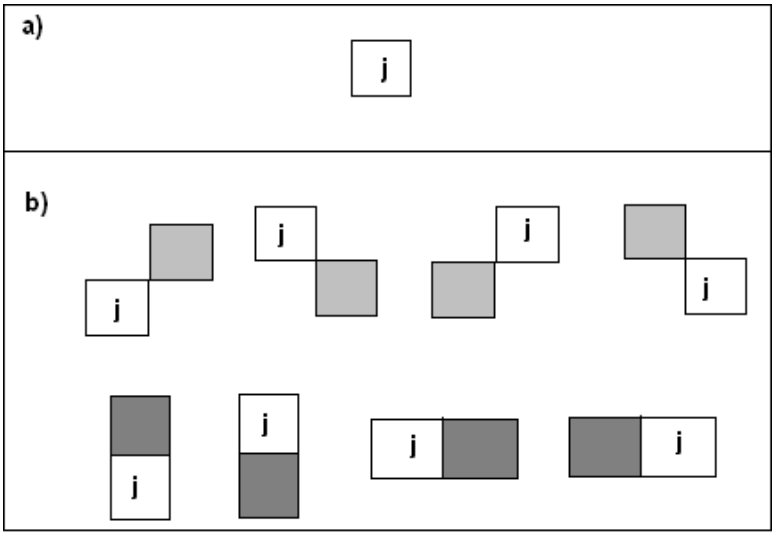


Figure 4.3: The Cliques a)C_1, is single cite clique b)C_2, is the pair-site clique.

Z , *is the normalizing constant given by equation 4.2*

$$Z = \sum_{x \epsilon X} e^{-\frac{1}{T}U(x)} \tag{4.2}$$

where, T, *is constant, termed as temperature*.

U(x),*is the energy function*

The energy function can be stated as sum of clique potentials over all possible cliques, given by :

$$U(x) = \sum_{c \epsilon C} V_c(x) \tag{4.3}$$

Cliques which are central to the GRF-MRF equivalence, can be stated as set of sites in which all the sites are neighbours of each other (figure 4.3). They can be visualized as subset of complete neighbourhood system. Hence, re-writing equation 4.3.

$$U(X) = \sum_{j \epsilon C_1} V_1(x_j) + \sum_{j,j' \epsilon C_2} V_2(x_j, x_{j'}) + \sum_{j,j',j'' \epsilon C_3} V_3(x_j, x_{j'}, x_{j''}) + ... \tag{4.4}$$

Now, Using equation 4.1, equation 4.3 and GRF-MRF equivalence (Theorem 4.1) the joint probability of MRF can be specified in terms of energy function and the clique potential, as :-

$$P(X) = Z^{-1} \times e^{-\frac{1}{T}\sum_{c \epsilon C} V_c(X)} \tag{4.5}$$

This helps to reduce the problem from global context to local context, hence, reducing the complexity and making it mathematically tractable.

4.1.4 Prior Energy

Using equation 4.5, *prior knowledge* about the image is determined. And the energy i.e. U(x) corresponding to it is called the prior energy. In case of image classification,

smoothness assumption is usually used as the prior information[8]. Various functions in the form of models exist in literature, which can be used to model the prior information and are known as smoothness priors. Some of those are *Ising Model, Auto and Auto-Logistic Model, Multi-Level Logistic Model etc.* [21, 20, 16, 8].

4.2 Discontintuity Adaptive Priors

Smoothness assumption, as discussed section 4.1.4 (previous section), implies that there are no abrupt changes in physical properties of the system i.e the DN values of the pixels do not change abruptly and are in coherence with each other. But again, this is just an assumption. But in practical scenario, the images are piecewise discontinuous [8] i.e. the discontinuity or abrupt changes is a reality. And if prior model is formed using this assumption of smoothness it would lead to over-smoothing and thus, undesirable results [23]. Hence, in this thesis, Discontinuity-Adaptive Smoothness priors and regularizers have been used to encode the prior energy.

The regularizer used in MRF are analytical regularizers, that are included in the prior constraints and penalize the irregularities. The aim of a MAP-MRF framework is to maximize the posterior probability (discussed in coming section) i.e to minimize the corresponding posterior energy.

The general form of these regularizers is given by the equation [23]:

$$U(x) = \sum_{n=1}^{N} U_n(x) = \sum_{n=1}^{N} \lambda_n \int_a^b g((x^n(y)))dy. \tag{4.6}$$

where, U(x) is the prior energy and also, called n_th order regularizer, where N is the highest order to be considered and $\lambda_n \geq 0$ is a weighing factor.

Here, the potential function is $g(x^{(n)}(y))$ is the penalty function and is analogues to clique potential discussed in section 4.1.3. And this function is what varies among different regularizers.

In this thesis we have used two Regularizers :-

1. **Standard Regularization**

   The standard regularizers use the quadratic function, given as :

   $$g_q(\alpha) = \alpha^2 \tag{4.7}$$

   The equation implies that higher irregularity in $x^{(n-1)}(y)$ at a site x, will lead to higher value of $|x^{(n)}|$ i.e. the derivative magnitude. This in turn leads to higher value of $g(x^{(n)})$, which causes increase in energy i.e. $U(x)$.

   This standard regularization method has been used previously by [13] with FCM based classifier. The shortcomings of this regularization method is [23]:

   (a) It considers a constant interaction, between neighbourhood sites, throughout.

   (b) and the smoothening strength becomes proportional to the derivative magnitude, $x^{(n)}$.

   According to [23], this will lead to over-smoothening even at discontinuities.

2. **Discontinuity Adaptive MRF Model or DA Model** The various regularizers differ only by the method they follow to check the interactions among the neighbouring sites and accordingly, adjust the strength. To avoid over-smoothing it is important that at the discontinuities the interactions must be managed accordingly and consequently, diminish. This is need is catered by DA models[21, 23].

There are four choices of potential function i.e. $g(\alpha)$ available. These are also known as *Adaptive Potential Function* or APF. The derivative of the APF is expressed as

$$g'(\alpha) = 2\alpha h(\alpha) \tag{4.8}$$

where, h is known as interaction function and influences the interaction between the neighbourhood sites. It is parameterized by γ and is written as h_γ and is also, termed as *Adaptive Interaction Function* (AIF).

And the strength with which a regularizer will perform smoothening is given by

$$|g'(x')| = |2x'h(x')| \tag{4.9}$$

where, α (equation 4.7) is given as $\alpha = x'(y)$.

All of these equations and DA models are the result of analysis of Euler equation, which is given as under :

$$u_x(x, x') - \frac{d}{dy}u_{x'}(x, x') = 0 \tag{4.10}$$

How DA priors works?

The necessary condition for a regularizer to be "discontinuity adaptive"is [23]:

$$\lim_{\alpha\to\infty} |g'(\alpha)| = \lim_{\alpha\to\infty} |2\alpha h(\alpha)| = C \tag{4.11}$$

where, C is a constant and $C\epsilon[0, \infty]$.

$C = 0$ implies zero smoothing strength i.e no smoothening at all and this will happen (as per equation 4.11) when $\alpha \to \infty$ and when $C > 0$ there will be bounding smoothening [21]. This also means, $h(\alpha)$ should vary inversely with α. In other words, at large $|\alpha|$, $h(\alpha)$ will be small and as $|\alpha| \to \infty$ then , $h(\alpha) \to 0$. And this is how, the DA model works.

Which DA model was used?

Out of four available DA models, given in [21], the DA model with following APF and corresponding AIF, respectively was selected.

$$g_{4\gamma}(\alpha) = \gamma(\alpha) - \gamma^2 \ln(1 + \frac{|\alpha|}{\gamma}). \tag{4.12}$$

$$h_{4\gamma}(\alpha) = \frac{1}{1 + \frac{|\alpha|}{\gamma}} \tag{4.13}$$

According to literature [21], this AIF (i.e. the fourth AIF among all the given models) allows smoothing at discontinuities but it is limited. This is overcomes the problem of boundless smoothing scene in case of standard regularization when $|\alpha| \to \infty$ and also, doesn't have abrupt end of smoothing as scene in case of other regularizer model such as Line Process Model (LP). This is because this APF's and AIF's band of convexity of B_γ is $B_{4\gamma} = (-\infty, +\infty)$.

Prior energy model based on both of the above stated, standard regularization as well as DA, were used seperately in this thesis. As well as, their results were compared. The details of which will be discussed in the ensuing chapters.

4.3 MAP-MRF Framework

Maximum A Posterior Probability or MAP with MRF forms a MAP-MRF framework. Here, the MAP-MRF model looks out for the most probable configuration for the sites in the image. This is done by maximizing the posterior probability $P(x|d)$ i.e. finding the MAP estimate.

$$x = \arg max_x P(x|d) \tag{4.14}$$

The Bayesian formula gives the relationship between posterior, conditional and prior probability:

$$P(x|d) = \frac{P(d|x)P(x)}{P(d)} \tag{4.15}$$

Following equation 4.1(seciton 4.1.2), the MAP estimate can be given as :

$$\hat{x} = \arg min_x U(x|d) \tag{4.16}$$

Using equation 4.15, equation 4.16 can be re-written as :

$$\hat{x} = \arg min_x (U(d|x) + U(x)) \tag{4.17}$$

where,

$\hat{x}$, *is the optimal variable value assigned to a site and in case of classification it can be stated as class label.*

$U(d|x)$, *is the conditional energy*

$U(x)$, *is the prior energy*

In nutshell, the objective is to find the appropriate objective function which includes this conditional and prior energy function. And to minimize the objective function so as to obtain the optimal variable for the site

4.4 Summary

In this chapter, the theoretical details about MRF has been discussed along with description of the prior models used in this thesis. In last section, MAP-MRF framework has been discussed which shows the relation between prior, posterior and conditional energy. It is important to understand and remember this MAP-MRF framework as it has been used in the classification developed in this thesis and explained further in chapter 6.

Chapter 5

Study Area and Data Preparation

This chapter has been dedicated to the remotely sensed data used in this research. It also includes the data preparation i.e. pre-processing of images, which is the first step of the methodology (See flow the diagram in figure 1.3, section 1.6).

The structure of the chapter is as follows. Section 5.1 describes the area from which the images have been taken to apply the classification method developed. Section 5.2 is about the sensors and the satellite(s), images of which have been used. In section 5.3, the discussion is on the pre-processing of the images and generation of reference data set. Section 5.4 summarizes the complete chapter.

5.1 Study Area

The images used in this thesis are of Sitarganj Tehsil, Udham Singh Nagar District, Uttarkhand, India. The state, Uttarkhand is in northern part of India, whereas Sitarganj is located in the southern part of the state. It is near PantNagar Agricultural University, famous for its participation during Green Revolution of India. In terms of Geographic lat/long, the area extends from 28°52'29"N to 28°54'20"N and 79°34'25"E to 79°36'34"E. The area consists of agricultural farms with sugarcane and paddy as one of the few major crops. Also, it has two reservoirs namely, Dhora Reservoir and Bhagul reservoir (can be seen as two huge water bodies, in the north and south-east part of the image shown in figure 5.1, respectively).

5.2 About the sensors

The images for this research thesis were taken from three different sensors all belonging to the same satellite : IRS-P6 (Indian Remote sensing Satellite) also known as, Resourcesat-1. This facilitates the use of the data set for classification as well as for referencing (s.a. AWIFS for classification & LISS III and LISS IV for referencing) providing the same illumination conditions. The sensors are, namely, LISS IV (Linear Self Scanning Sensor), LISS III and AWIFS (Advanced Wide Field Sensor). Table 5.1 has the relevant details of these sensors.

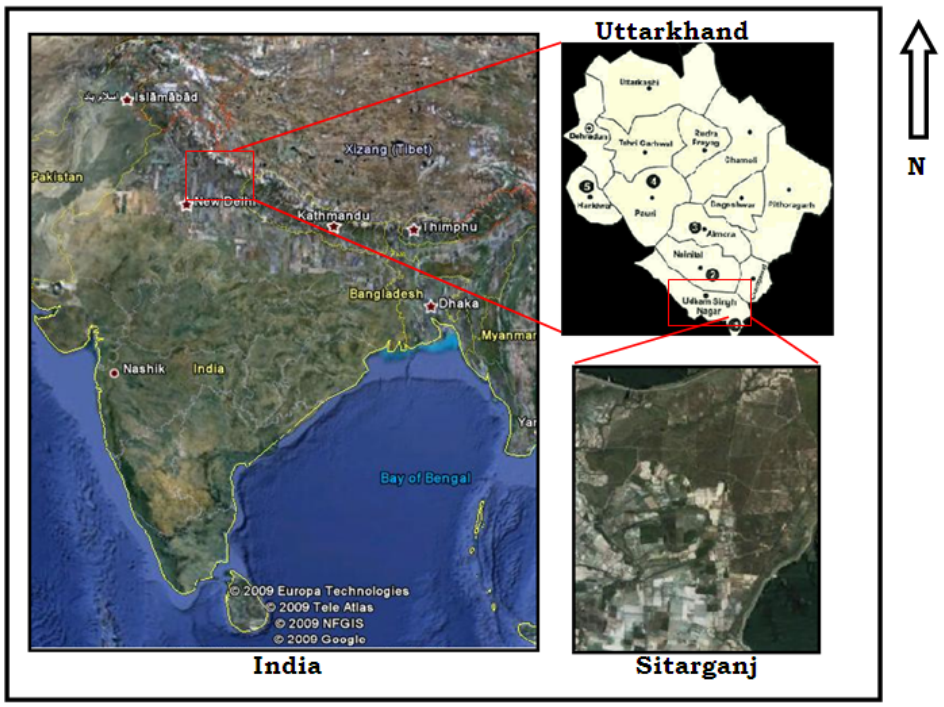


Figure 5.1: Location of area under study (*Source:*GoogleEarth,accessed on :13 Nov 2009)

5.3 Data Preparation

This section covers relevant details about image pre-processing and the generation of reference data set.

5.3.1 Geometric correction of LISS IV images

First step in image pre-processing was image-to-map rectification of LISS IV image with Survey of India (SOI) toposheet. The LISS IV image (dated : 15 October 2007) was geo-registered precisely with SOI toposheet, numbered $53\frac{P}{9}$.UTM projection was used with spheroid and datum being Everest North, in zone 44. For this purpose the SOI toposheet was scanned and converted into digital form before geo-registration.

Table 5.1: IRS-P6's sensors' specification

Specifications	**LISS IV**	**LISS III**	**AWIFS**
Spatial Resolution(m)	5.8	23.5	56
Swath(km)	23.9(MX Mode) 70.3(PAN Mode)	141	740
Spectral Bands(microns)	0.52 - 0.59 0.62 - 0.68 0.77 - 0.86	0.52 - 0.59 0.62 - 0.68 0.77 - 0.86 1.55 - 1.70	0.52 - 0.59 0.62 - 0.68 0.77 - 0.86 1.55 - 1.70

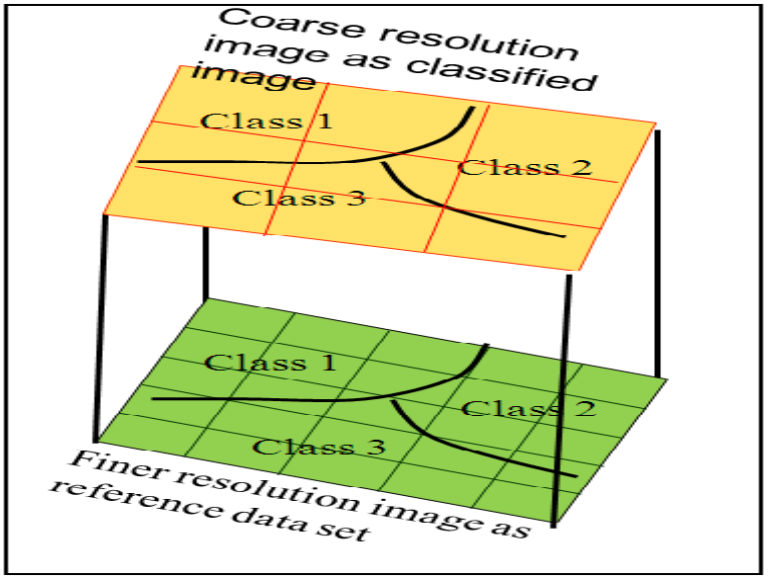


Figure 5.2: Comparing coarser resolution image with finer resolution image

LISS IV image was resampled at 5m resolution. The resampling was important in particular, for accuracy assessment as each finer resolution image was used as a reference data set for classified images of coarser data set. The resampling of all three images i.e. LISS IV, LISS III and AWIFS was done in such a way that the pixel size in all three images formed a ratio of 1:4:12, respectively. Thus, resampled size of pixel in LISS IV is 5m.

Resampling is needed for geometric transformation and is similar to convolving the input image with a uniform weighing function. The sampling method used here is **Nearest Neighbour Sampling**, where the resampled pixel takes the value of the nearest pixel in the neighbourhood. It is a fast resampling method, which do not require calculations. But the implications of this resampling is introduction of geometric discontinuities in the order of plus-minus half pixel [26].

5.3.2 Geometric Correction of LISS III image

Geo-registration of LISS III image was done with geometrically corrected LISS IV image with same specifications of datum, spheroid and projection as that of LISS IV.

The image was resampled to the spatial resolution of 20 m, as per already discussed during the geometric correction of LISS IV.

5.3.3 Geometric Correction of AWIFS

Similar to that of LISS III, geo-registration of AWIFS also done with LISS IV in same projection, at same spheroid and datum as that of LISS IV and LISS III.

The resampling of AWIFS, using nearest neighbourhood, to spatial resolution of 60 m was done.

5.3.4 Generation of Reference data set

In this thesis, the comparison of the classified data set has been made with reference data set generated by classifying the finer resolution images. Here, the reference data

for AWIFS was LISS III and LISS IV. Similarly, the reference data set for LISS III was obtained from LISS IV.

There has been an emphasis on generating the reference data set from a finer resolution image, instead of using data from field because of the following reasons :-

1. It is a difficult task to locate within-pixel classes on the ground.
2. The certainty to which the classes can be identified on the ground is a subjective issue.[25, 5]
3. Ground data generated itself is a kind of classification and thus, may have errors. [25]
4. In-accessibility of the study area is another hindrance in obtaining the ground data.

When the ground truth is not available as stated above, aerial photography is another option [25]. But aerial photography is not possible in this case because of the expenses incurred, government regulations and also, the satellite images are dated 2 years back. Thus, finer resolution images have been used.

In case of satellite image as reference data set, it is preferable to use both test and reference images having same date of acquisition to avoid discrepancies cropping in due to different illumination conditions. That has been taken care in this research, as all the three images i.e. reference data and test data are taken from the same satellite (section 5.2) and are of same date.

To compare the two reference data set with the test data, it is important to classify the reference data set as well. For classification of the finer resolution to be used as reference data set, there are two methods :-

Method 1: Multiple Resolution

This method implies aggregating the finer resolution pixels to form a coarser resolution pixel. In the present case, it is to combine 4×4 i.e. 16 pixels of LISS IV to form the coarser resolution pixel for the LISS III. Similarly, for AWIFS, $12 \times 12 = 144$ pixels of LISS IV will be aggregated to form one pixel of AWIFS. And then classify it.

Though, visually it seems to provide the required outcome. But on closely examining it, the resultant membership value is a combination of proportion of the pixels that constitute the class. In other words, they are not the result of ambiguity or uncertainty in the classes. This also implies, that the sum of membership value of a pixel should be one [46].

Also, it assumes that the pixels from finer resolution data set are pure which is not true. A pixel at finer resolution (5m in the case of LISS IV) do have mixed pixles and sub-pixel objects[5, 3].

Hence, this method is not suitable for assessing the accuracy of the method proposed in this thesis.

Method 2: Using sub-pixel Classifier

Second method is to use a sub-pixel classifier to generate the reference data set.

Advantages

- This method has advantage over method 1, as the membership values generated will be the result of uncertainty or ambiguity of classes and may not sum to one (unless, the constraint is otherwise specified as in FCM).

- Also, from the preceding statement it is evident that the pixels are not assumed to be pure, hence, giving more realistic results.

Hence, in this thesis, method 2 has been selected for generating the reference data set.

Disadvantages But as all the methods have their own advantages and disadvantages. This method has two issues of concern :

1. The resolution of classified data set are different. To compare the reference and classified data the reference data set is either resampled or aggregated mean of pixels in reference data set is taken. In both the cases, there is an unavoidable error is bound to be present which can be attributed to difference in scale. (For resampling, implications have been discuses in section 5.3.1)
2. When the classifier for the reference data set is different from test data there are two cases that may arise. First, the performance of classifier data set is better in terms of accuracy than the classifier that is being tested. Second, the classifier that is being tested has better performance. In both the cases absolute measure of performance cannot be given and thus, can be stated that accuracy assessment will be of dubious validity.

In this research there was a choice two classifiers for generating the reference data set, as enumerated:-

1. **Using Support Vector Machine (SVM) classifier**.
   SVM is a statistical learning algorithm and has been shown to be an efficient classifier, especially, for finer resolution data sets [28]. Also, the other reason for selecting SVM is that, it has proved to perform better in terms of accuracy as compared to FCM and PCM [28, 30]. But using this classifier causes error both because of difference in resolution and due to different classifier (as per points stated in disadvantages).
2. **Using the PCM-MRF classifier or same classifier**
   The other method is to use the same classifier for both test and reference data set. As in this case, te classifier is PCM-MRF , which is another option to be used for classifying reference data set as well. This will eliminate at least one drawback i.e. caused by use of different classifiers. And error will be only because of difference in resolution of the images.

In this research, both the classifiers were used to generate the reference data set and the results of accuracy assessment are given in chapter 7 (section 7.1.1 and section 7.1.2), followed by discussion.

5.4 Summary

In this chapter, the data used for this research has been mentioned, including the necessary pre-processing done. Apart from the technical benefits of having images of same date and from same sensor, the another reason to select this particular image set is its importance in the Indian context.

RESOURCESAT-I (IRS-P6) is tenth in the series of Indian Remote-sensing Satellites (IRS) with the objective of natural resource management. With 24 day repeat cycle, the sensors AWIFS, LISS III and LISS IV, the objective is to provide large spectral and spatial coverage of natural resources in form of remote sensing images. The

vast amount of data captured (in the form of digital images) is being used in preparing thematic maps, which are further analysed and thus, utilized in study of crop yield, crop stress, disease surveillance, disaster management etc [27]. This requires more and more accurate classification algorithms which can classify the images with near to real-life scenario.

There is need for further pre-processing before the images can be used for classification that includes geo-referencing and image registration. It also includes generation of reference data set, which in this research, has been generated from the finer resolution images instead of having ground truth. The error induced during these processes are inevitable and carried along in the further steps including classification. Thus, error/inaccuracy in the classifiers performance is not completely due to the classifier itself, but also due to these pre-processing steps.

Chapter 6

A Hybrid Approach to Image Classification : *PCM-MRF*

"Thought and theory must precede all salutary action; yet action is nobler in itself that either thought or theory" - William Wordsworth

This chapter describes the "action" i.e. the steps that followed the development of PCM-MRF based classification approach in the thesis.

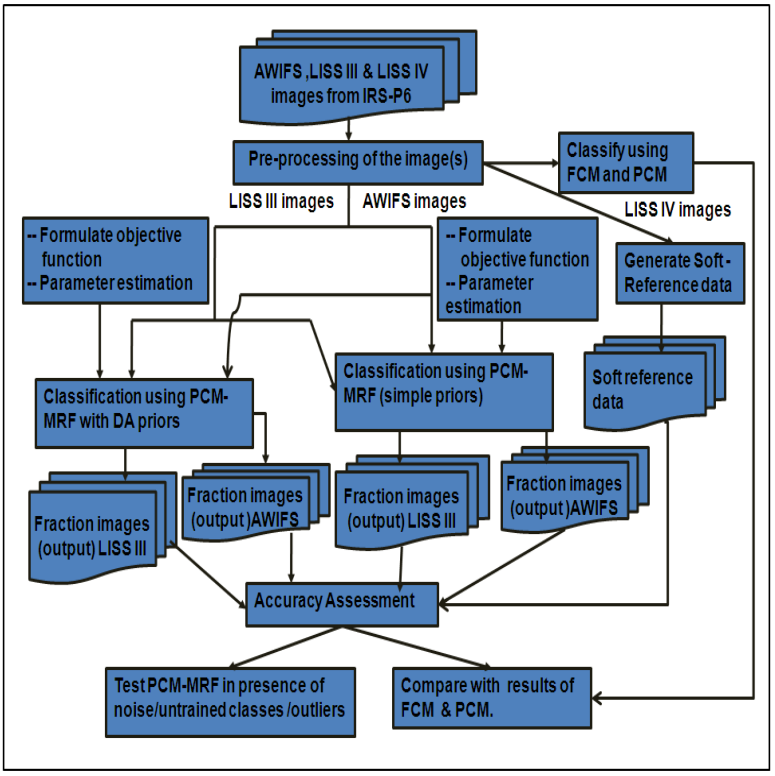


Figure 6.1: The flow chart of the steps in methodology

The chapter is structured into sections. An attempt has been made to make each section consistent with the individual steps, as given in the Methodology flow chart (figure 6.1).

Note: ***Step 1 & 2*** *i.e. "Acquire images from IRS-P6" and "pre-process the images" has been covered in chapter 5.*

6.1 Formulating the Objective Function

The procedure for solving a problem using MAP-MRF framework calls for the formulation of objective function first. Here, two objective functions have been formulated and utilized, for the reasons already stated in section 4.2, chapter 4. Both of them have some similarity to the objective function of PCM and differ on the aspect how the neighbourhood information has been included. The sequence of discussion starts with the objective function of PCM, followed by two sub-sections, each corresponding to one of the two objective functions.

The objective function of PCM discussed in equation 3.3, chapter 3, is minimized with respect to membership function u_{ij} and the mean of the cluster v_{ij}, used in calculating the distance d_{ij} given in equation 3.3 and restated here

$$J(L,U) = \sum_{i=1}^{C}\sum_{j=1}^{N}(u_{ij})^m d_{ij}^2 + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-u_{ij})^m. \tag{6.1}$$

Equation 6.1 includes the information about the distance of the feature vector (that forms the pixel) from the cluster mean in the feature space but it does not include information on spatial context. The spatial context here includes, as discussed, the influence of the neighbouring pixel on the target pixel in the image space.

As discussed in section 4.3, chapter 4, the MAP-MRF framework works by maximizing the posterior probability which is related to prior and conditional energy (equation 4.15 to equation 4.17). The global posterior energy has been defined in equation 6.2 (given below) :-

$$U(x|d) = U(d|x) + U(x) \tag{6.2}$$

Where, x is the class label that minimizes the global posterior energy. The conditional energy function $(U(d|x))$ comes from the spectral information and the prior energy function $(U(x))$ is the smoothness prior or the spatial information.

In order to balance the contribution of the two information i.e. the spectral and spatial information a parameter λ is introduced in equation 6.3 .

$$U(x|d) = (1-\lambda)U(d|x) + \lambda U(x) \tag{6.3}$$

The value of lambda may vary from 0 to 100, but for the sake of simplicity lambda has been limited between 0 to 1 in this thesis. And as the equation 6.3 suggests, higher the value of λ the contribution of smoothness prior increases and that of spectral information decreases and vice-versa.

6.1.1 Objective function with Standard Regularizer

Equation 6.4 is the first objective function formulated with standard regularizer. This objective function is referred to as **PCM-MRF I** from now on.

$$U(u_{ij}|d) = (1-\lambda)[\sum_{i=1}^{C}\sum_{j=1}^{N}(u_{ij})^m d_{ij}^2 + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-u_{ij})^m] + \lambda[\sum_{j=1}^{N}\sum_{i=1}^{C}\sum_{j'\epsilon\mathcal{N}_j}\beta(u_{ij}-u_{ij'})^2] \tag{6.4}$$

Where, the symbols have the meaning as explained in Chapter 3 and Chapter 4. Except, β, which is the weight given to the neighbouring pixels in the window. And x has been replaced with the membership function u_{ij}

In equation 6.4, spectral information has been included using the objective function of PCM (equation 3.3) and spatial information is incorporated in the form of Standard Regularizer.

The similarity between PCM and MAP-MRF lies in the goal to minimize the objective function and the same has been exploited.

6.1.2 Objective function with Discontinuity Adaptive Priors

As discussed in section 4.2, chapter 4, the Discontinuity Adaptive Priors (or DA priors) avoid over smoothing. They have been used to preserve the edges as they account for discontinuities (e.g. edges of the agricultural fields). Edge preservation is important because the errors usually occur around edges and thus, contribute to inaccuracy in overall accuracy assessment [32]. Also, if the edges are preserved

The objective function with DA priors is (equation 6.5):-

$$U(u_{ij}|d) = (\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m} D_{ij}^{2} + \sum_{i=1}^{C} \nu_i \sum_{j=1}^{N} (1-u_{ij})^{m} \times)(1-\lambda) + \lambda(\sum_{j=1}^{N}\sum_{i=1}^{C}\sum_{j' \epsilon N_j} (\gamma|\eta| - \gamma^2 \ln(1+\frac{|\eta|}{\gamma}))) \quad (6.5)$$

Where, all the symbols have the usual meaning, except two new symbols introduced.

γ, *AIF (see section 4.2, chapter 4), is parameterized by* γ

η, *is* $u_{ij} - u_{ij'}$ *i.e. difference between target pixel's membership value and it's neighbourhood pixel's membership value in a neighbourhood window*.

Equation 6.5 uses the APF discussed in section 4.2, chapter 4 and replaces the standard regularizer with the given APF. From now on, equation 6.5 i.e. the second objective function is referred to as **PCM-MRF II**.

Equation 6.4 and 6.5 are the final objective functions that give the posterior energy. In succeeding sections, emphasis is on utilization of these function for classification.

6.2 Simulated Annealing and Sampling Algorithm

Once, the objective function, which corresponds to the posterior energy is formulated the task is to determine the optimal solution. Optimal solution here is the membership values of the pixels which will minimize the posterior energy. The minimization of energy function involves finding the global minima.

Annealing is one of the popular methods, which determine the global minima without getting stuck at the local minima [10, 17, 21]. Annealing can be broadly categorized into two categories [21]:-

1. Deterministic, such as graduated non-convexity algorithm (GNC)
2. Stochastic, such as Simulated Annealing.

GNC and Simulated Annealing are two popular algorithms in case of MRF related problems [21]. Simulated Annealing (proposed by [37] & [38] independently) was used in this thesis. It has been found efficient in finding the global minimum of the objective function which may have various local minima as in case of image processing with MRF. Thus, has been extensively used in solving problems related to image processing such

as super-resolution mapping using MRF[49, 54], texture based image segmentation using MRF [33] and SAR image segmentation [36], image classificationDuttaThesis, and vision after the classic paper by Geman & Geman in 1984.

As the name itself suggests, "Simulated Annealing"emulates the physical annealing process where the solid is first melted by heating it to a very high temperature and then, it is slowly and gradually cooled down to the desired "frozen"state which has the minimum energy configuration[39]. In the context of image classification, it implies first perturbing the labels of the pixels randomly which is similar to melting of the metal. Then image is passed through the cooling schedule (described in equation 6.7) and by determining the global minima the final classified image is obtained i.e. the "frozen state".

Psuedo Code *After deciding upon the neighbourhood system and estimating the parameter of the objective function (this will be dealt with in the next section, section 6.3) following steps have been followed for simulated annealing [21, 8].*

Begin

1. Set initial state of image which is the image classified using PCM.
2. Set initial temperature T=4 and no. of iterations N=1000.
3. Set the cooling schedule i.e. $T(t) = \kappa T(t-1)$[37], where, $T(t)$ is the temperature at time t and κ is the update rate of temperature.
4. Estimate the new membership values using Gibbs Sampler.
5. Calculate the initial posterior energy using the objective function (either equation 6.4 or 6.5) $U(x|d)_1$ and the new posterior energy $U(x|d)_2$.
6. Find $\Delta = U(x|d)_1 - U(x|d)_2$,
   if $\Delta > 0$,

   replace previous membership values with new membership values

   else if $exp(\Delta) \geq random[0, 1]$

   replace previous membership values with new membership values.
7. Repeat step 4 to 6 N times.
8. When no. of iterations = N+1, update T using the cooling schedule described in step 3.
9. Repeat step 4 to 7 until $T \longrightarrow 0$ i.e. "frozen"state is reached.

end

In step 2, T i.e. intial temperature was estimated as 4 (discussed in section 7.1.2). Equation 6.6 below, gives a popular cooling schedule [21].

$$T(t) = \frac{T_o}{log(1+N)} \tag{6.6}$$

But this has been rendered too slow for practical purposes [21]. Thus, cooling schedule given by Kirkpatrick et al. in 1983 [21] has been used and described in equation 6.7 below (also, in step 3 of pseudo code):

$$T(t) = \kappa T(t-1) \tag{6.7}$$

Where, κ varies from 0.8 to 0.99 and has to be estimated based on the problem.

Sampling Algorithm : Gibbs Sampler As mentioned in step 4 of the pseudo code, Gibbs Sampler has been used to generate the next membership value. A Gibbs Sampler is a special case of Metropolis-Hastings algorithm for sampling, where, based on the conditional probability the next configuration is generated [40]. As expressed in equation 6.8 below :

$$P(x(t)|x_{\mathcal{N}}) = \frac{e^{-U(x)/T}}{\sum_x e^{-U(x)/T}} \tag{6.8}$$

This produces a Markov chain and comes under the umbrella of Markov Chain Monte Carlo Methods (MCMC)[40].

6.3 Parameters to be estimated

Simulated Annealing process described in the pseudo code (section 6.2), requires following parameters as input.

1. Fuzzifier (m)
2. Initial Temperature (T_o)
3. Temperature Update Rate (k)
4. Lambda (λ)
5. Gamma (γ)

These parameters do not have fix values and estimation is required. According to [21], there is no standard technique to estimate these parameters. Various methods such as comparing RMSE, Total Energy etc. has been used in estimating these parameters [13, 49]. In this work, estimation of parameters has been done using RMSE, which gives the relative measure (explained in section 6.4.2). In order to have the absolute measure of uncertainty, entropy has been used (explained in section 6.4.4).

The results of parameter estimation has been discussed in chapter 7 (Results and Discussion).

6.4 Methods adopted for Accuracy Assessment

After classifying the images and obtaining the classified output it becomes is imperative to measure its accuracy. This aids in quantitatively assessing the efficiency of the classification method in terms of its accuracy. In this section, techniques used for accuracy assessment in this research work, has been described.

6.4.1 Fuzzy Error Matrix and other operators

For assessing the results of classified outputs of mixed pixels, there is has not been any standard procedure yet. Sometimes, the output is defuzzified to produce hard outputs but it is the loss of information and also, the purpose of classifying the mixed pixels is nullified.

Therefore, to preserve the result of sub-pixel classified outputs various methods have been proposed. E.g. : Fuzzy Error Matrix (FERM) [53], composite operators MIN-MIN, MIN-LEAST and MIN-PROD [46] and Sub-pixel confusion-uncertainty matrix (SCM) [43].

1. **FERM**

   Fuzzy Error Matrix by [42] is based on the traditional error matrix but unlike traditional error matrix which requires hard classified images as input, it takes the fraction images as the input. Thus, rows and columns have real values instead of the integers. It is based on MIN operator (given as equation 6.9 below) which gives the complete overlap of the fuzzy sets and also, known as intersection operator[43].

   $$Pn_{ij} = MIN(Pn_i, Pn_j) \tag{6.9}$$

2. **Composite Operators and SCM** The composite operators given by [46] especially, **MIN-LEAST** and **MIN-PROD** are suitable for the multiple-resolution based classification and thus, follow the constraint i.e. $\sum_{i=1}^{C} u_{ij} = 1$ (see, equation 1.1, chapter 1). Thus, have not been found suitable for the accuracy assessment of the possibilistic based method used in this thesis. The possibilistic based methods gives the membership value acknowledging the ambiguity, unlike, that of multiple-resolution where each pixel has membership according to proportion contribution by that pixel at coarser resolution.

   SCM also follows the constraint that of MIN-PROD and MIN-LEAST (equation 1.1) making it unsuitable for the use in possibilistic based method.

3. **MIN-MIN**

   MIN-MIN is a composite operator which uses MIN operator for calculating both diagonal (agreement) and off-diagonal elements(disagreement) of the error matrix[43].

For accuracy assessment of the classified output of the PCM-MRF method(s)(developed in this research), FERM and MIN-MIN has been found suitable. Where-as, MIN-PROD, MIN-LEAST and SCM are not suitable for assessing the accuracy PCM-MRF because of their hyperplane constraint.

6.4.2 Root Mean Square Error

Root Mean Square Error is given by taking a square root of the sum of squared difference between the membership values of the classified image and the reference image, see equation 6.10:-

$$RMSE = \frac{\sqrt{\sum_{j=1}^{N} \sum i = 1^{C} (C_{ij} - R_{ij})^2}}{M \times N} \tag{6.10}$$

Where,
C_{ij}, is the membership values from the classified image
R_{ij}, is the membership values from the reference image
$M \times N$, is the size of the image

RMSE gives the measure of both systematic and random errors [44]. It is the average measure of how much the membership values of the classified image differs from the membership values of the reference data set. The RMSE values are always greater than or equal to zero, as it is evident from the equation 6.10. To interpret RMSE, a good result is one, where RMSE values is minimum i.e. tends to zero.

For the given data, RMSE is calculated in two ways : Global and per class.

Global RMSE is the RMSE of the complete image i.e. all the fraction images and is given by equation 6.10. For calculating RMSE per class equation 6.11 has been taken into account.

$$RMSE(perclass) = \frac{\sqrt{\sum_{j=1}^{N} (C_{ij} - R_{ij})^2}}{M \times N} \tag{6.11}$$

6.4.3 Correlation Coefficient

Correlation Coefficient is used to measure the linear association between the two variables say X and Y. Among all the available correlation coefficients "Pearson-moment correlation coefficient" is best known [50].

In this thesis, the two variables of which the correlation is to be determined are the membership values of the classified image and the membership values of the reference image . And is given by equation 6.12 :

$$r = \frac{cov(R, C)}{\sigma_R \sigma_C} \tag{6.12}$$

Where,

$cov(R, C)$, represents the covariance between the Reference(R) and Classified (C) data.

σ_R & σ_C , are the standard deviation of R and C, respectively.

The range of r is from -1 to +1. If the variables (R and C) are in perfect straight line then, $r = +1$ implies increasing linear association and $r = -1$, is decreasing linear association. $r = 0$, which is a special case, shows no correlation between the variables. A value from 0.5 to 1, states a strong correlation between two variables [45].

6.4.4 Measure of Uncertainty : Entropy

FERM, MIN-MIN etc. are used to evaluate the performance of the classifier in terms of its correctness. Where as, RMSE and correlation coefficient (section 6.4.2 and 6.4.3) are the uncertainty measures. But these methods, are defined based on the difference between the expected and actual results and are relative measures. Thus, they are sensitive to error variations and not to the uncertainty variations [47].

According to [47], entropy is an absolute measure of uncertainty, calculated only from the classified data without requiring any other external information. Hence, in this research, entropy has been incorporated as another measure of uncertainty on experimental basis.

Based on Shanon's Entropy theorem and as discussed in [48, 47, **?**] following entropy based on fuzzy membership value can be calculated as given in equation 6.14 :-

$$H = -\sum_{i=1}^{c} u_{ij} \log_2(u_{ij}) \tag{6.13}$$

But as in the case of PCM, where the membership value does not follow the constraint of equation 1.1, the above given entropy theorem can be utilized by rescaling as in equation 6.15 [41].

Thus, the average entropy (based on Shanon's Entropy Theorem) of the complete image can been calculated using the following formula [47, 48, **?**]:

$$H_{avg} = -\frac{\sum_{i=1}^{C} u_{ij} log_2 u_{ij}}{\sum_{i=1}^{C} u_{ij}} \tag{6.14}$$

Also, [15] have used the above given entropy measure for validating the clusters formed during unsupervised clustering using FCM. In this research work, entropy has been taken as a measure to optimize 'm'. Also, it has been used in combination with other measures to optimize parameters related to MRF and it's effect has been observed (discussed in section 7.1).

6.5 Summary

In this chapter, a complete description of how the objective function was developed and used in classification has been given. The objective function using both standard regularizer and DA prior model was developed in section 6.1. The objective function was then mapped to the MAP-MRF framework described in section 4.4 and was minimized using Simulated Annealing algorithm discussed in section 6.3. It also, talks about the cooling schedule used and the steps followed in finding the global minima using Simulated Annealing. To assess the performance of classifier in terms of accuracy, various accuracy assessment measures and uncertainty measures, found suitable and thus, used in this thesis (quantitative results are given in chapter 7) were discussed in section 6.4. It was found that SCM, MIN-PROD and MIN-LEAST weren't suitable for assessing the accuracy of PCM based classifier, where the hyper-plane constraint is not followed.

Chapter 7

Results and Discussion

To understand and illustrate the efficiency (in terms of accuracy) of the PCM-MRF classification method developed in this thesis, it is imperative to perform the accuracy assessment and review the results. It is tough to pin-point a single accuracy assessment procedure that can establish the efficacy of a sub-pixel classifier. Therefore, as stated in section 6.4, chapter 6, various methods have been employed for the assessment of the performance of the developed classifier.

Section 7.1 consists of discussion on how the parameters (section 6.3) have been estimated and their estimated values have been given. Section 7.2 is on the performance of PCM-MRF with standard regularizer. Results of DA based PCM-MRF, with special emphasis on edges, have been discussed in section 7.3. The effect of untrained classes on PCM-MRF has been discussed in section 7.4. The results in section 7.2, section 7.3 and section 7.4 have been compared with the results of supervised PCM. Results of supervised FCM have also been generated and given in Appendix A. And at the end section 7.5 gives the overview of the chapter 7 in form of summary.

7.1 Estimating the Parameters

The parameters that were listed in section 6.3, chapter 6 were estimated which are given in following subsections.

7.1.1 Fuzzifier (m)

Fuzzifier or 'm' is a weighing exponent, where $m \in [1, \infty)$. If m = 1, then membership values are crisp i.e. when the point has distance $(d^2(x_j, \beta))$ greater than η_i they'll have zero membership to that cluster, particularly, in PCM [11]. Entropy has been used to determine the uncertainty[48, 47] and optimization for FCM and possibilistic clustering [15] as seen in the literature survey. Thus, in this research estimation of m was done by calculating entropy of the classified output. The estimated value of m for PCM-MRF was found to be 3.0, as given in Figure 7.1.

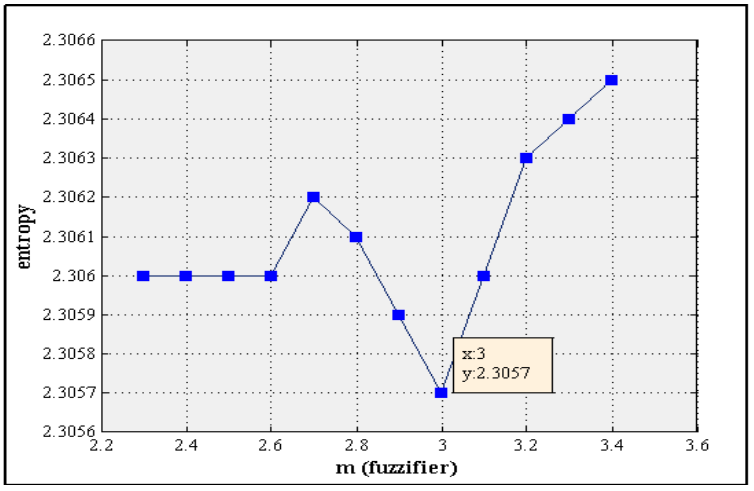


Figure 7.1: Estimating m

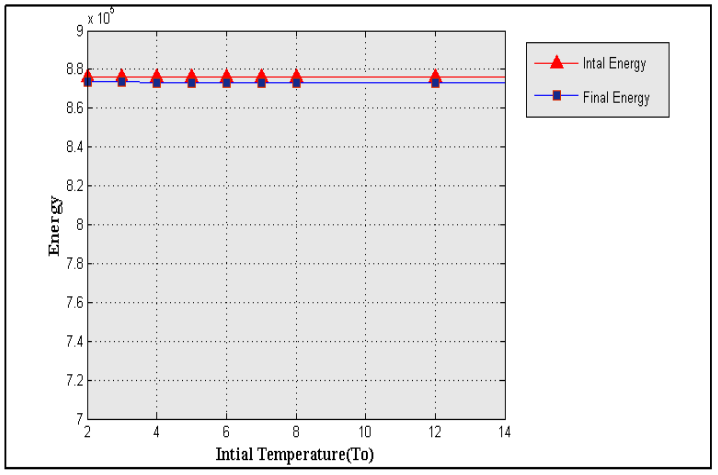


Figure 7.2: Estimating T_0

7.1.2 Intial Temperature (T_o)

According to [21, 10, 50, 49, 13], T_o is should be 3 or 4 for the purpose of image analysis. It can be justified with the results obtained by the experiments performed during this research work. As seen in figure 7.2, value of T_o above 4 did not show much improvement. Hence, for this research $T_o = 4$ was selected.

7.1.3 Temperature update rate (k)

The temperature update rate i.e. k determines the speed of cooling schedule (discussed in section 6.2, chapter 6). A low update rate states faster cooling and a high update rate implies slower cooling. The range of k is said to lie between 0.85 and 0.99 [21].

The appropriate value of k was determined by estimated by RMSE and entropy for both AWIFS and LISS III. The optimization is shown in Table 7.1 and Table 7.2. It is observed from the result obtained that update rate of LISS III ($k = 0.92$) is higher than that of AWIFS ($k = 0.9$).

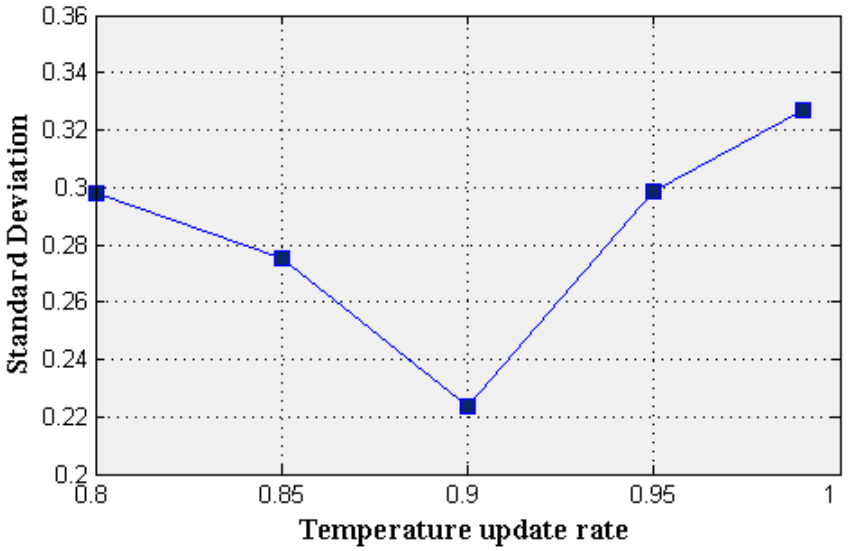


Figure 7.3: Reproducibility of PCM-MRF solution

Number of iterations were checked to observe the time taken by Simulated Annealing to converge. The convergence time depend on cooling schedule which includes T_o and k. For AWIFS, $k = 0.9$ and took 115 iterations on an average for convergence. Whereas for LISS III images, $k = 0.92$, the cooling process was slow and it took on an average 456 iterations. For both AWIFS and LISS III, large number of updates were observed towards the end of cooling schedule at each temperature value.

Also, reproducibility of PCM-MRF solution was observed at various value of k by keeping other parameters constant (Figure 7.3).

Table 7.1: Optimizing k for AWIFS a) Using RMSE b) Using Entropy)

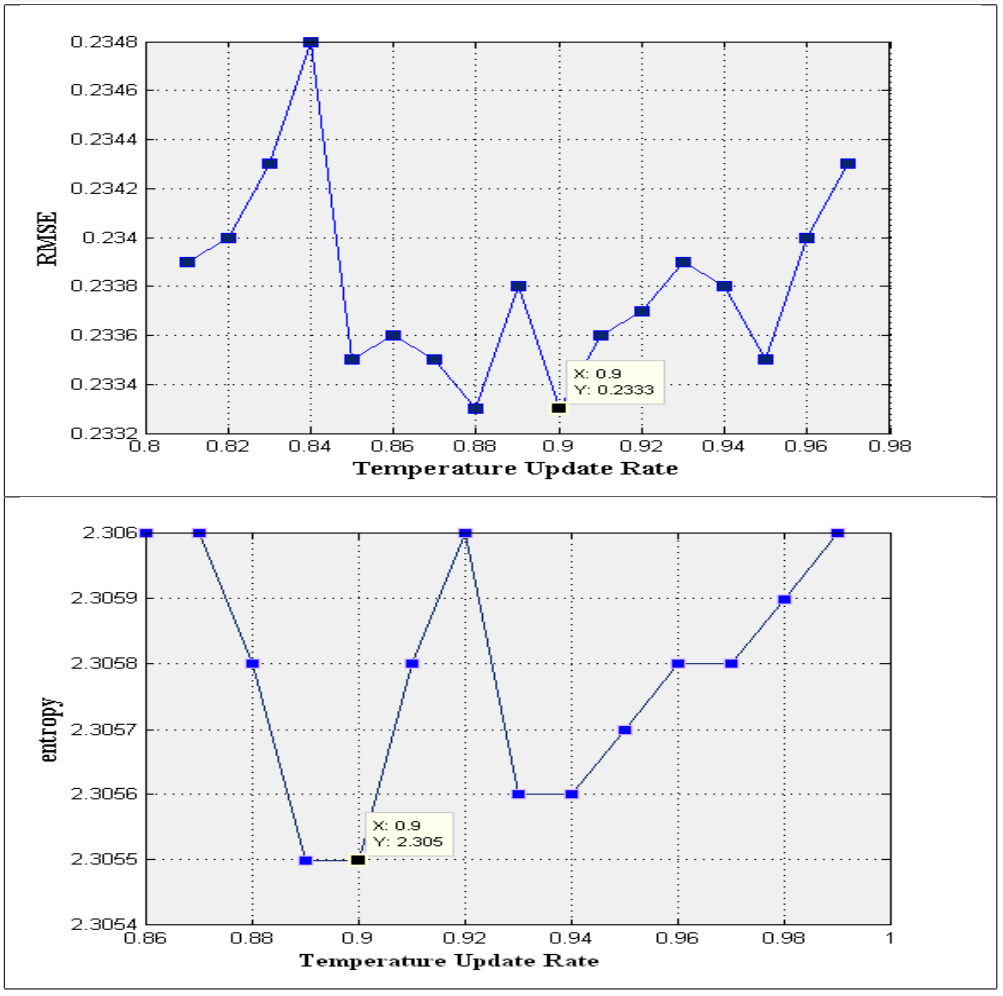
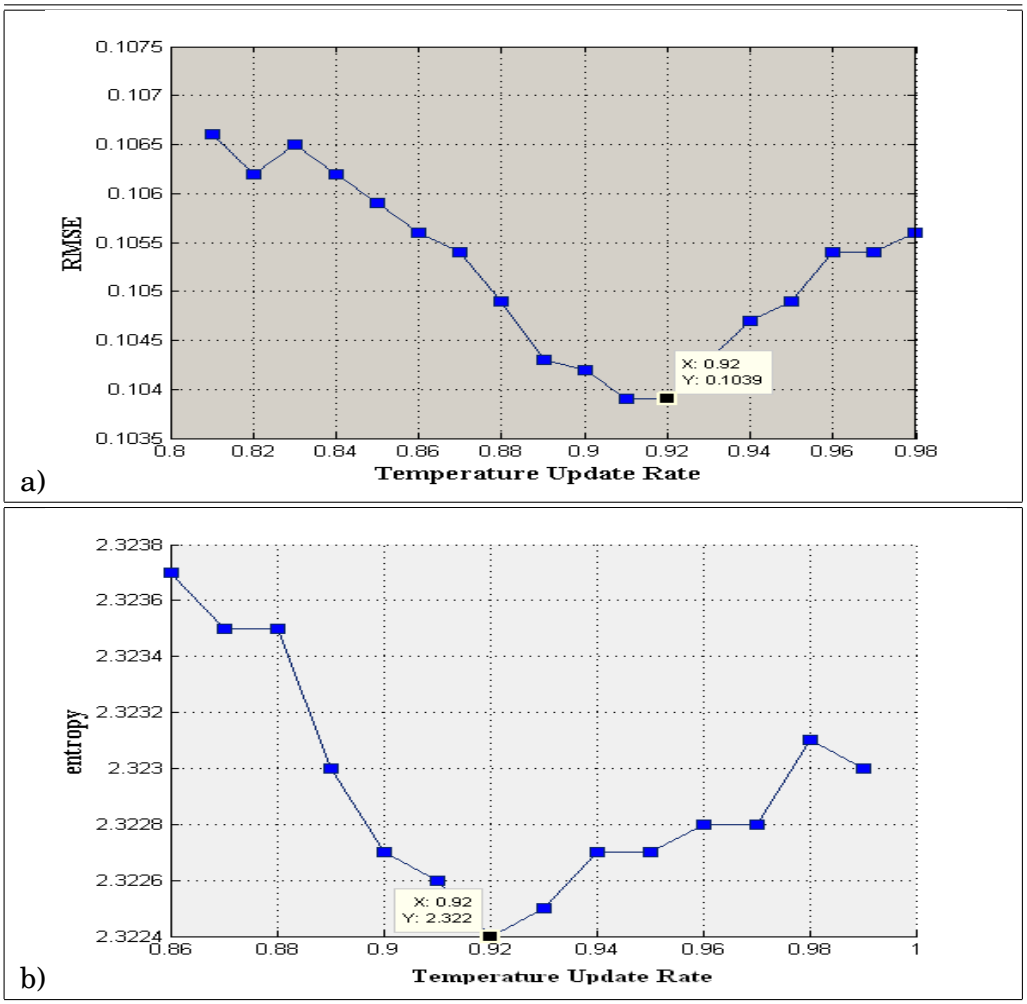


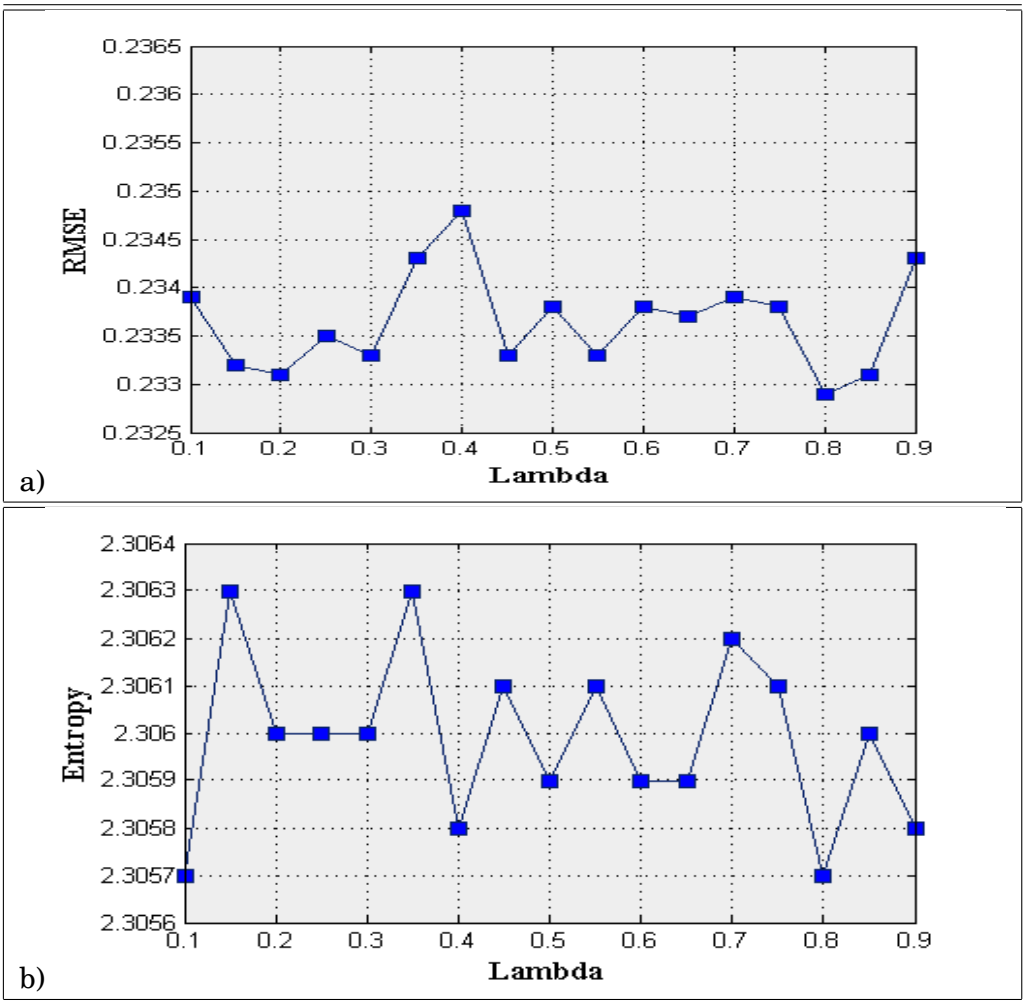
Table 7.2: Estimating k for LISS III: a) Using RMSE b) Using Entropy)



7.1.4 Lambda(λ)

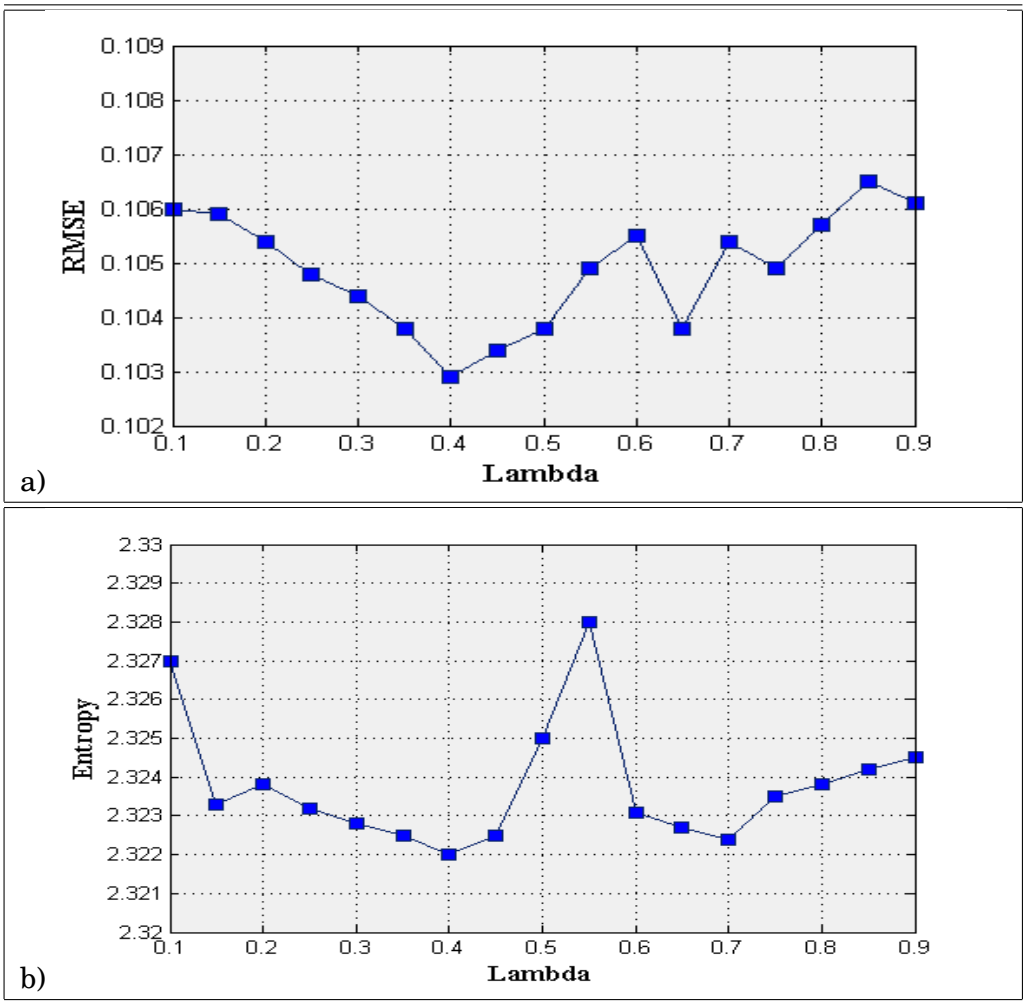
As mentioned in section 6.1, λ is the smoothness parameter that controls the contribution of spectral vs. spatial information. RMSE and entropy both were calculated for various values of λ. Table 7.3 and 7.4 gives the graph showing estimation of λ.

Table 7.3: Estimating λ for AWIFS: a) Using RMSE b) Using Entropy



It is found that the value of lamnda varies for AWIFS ($\lambda = 0.8$) and LISS III images($\lambda = 0.4$).

Table 7.4: Estimating λ for AWIFS: a) Using RMSE b) Using Entropy



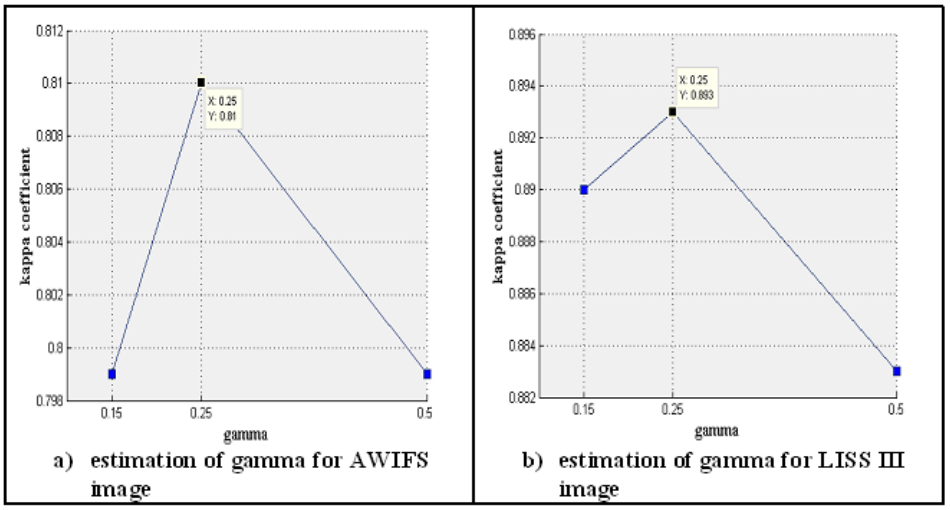


Figure 7.4: Estimating gamma

7.1.5 Gamma (γ)

In DA models, γ determines the rate at which AIF (refer section 4.2, chapter 4) approaches zero. It can be considered the threshold value after which interaction between two pixels should minimize, or in other words $|f_i - f_{i'}|$ i.e. the difference between target pixel value and the neighbouring pixel becomes very large [22].

Estimation of γ was done by referring to the fuzzy kappa coefficient(s). The data set was tested for three different values of γ viz. 0.5, 0.25, 0.15 and the results are given in the Figure 7.3.

7.2 Results of PCM-MRF with standard regularizer: PCM-MRF I

This research is aimed at performing level I classification of the land cover classes. Based on the visual interpretation from the LISS IV, LISS III and AWIFS images following land cover classes were selected:

1. Water
2. Agricultural field with crop
3. Dry agricultural field without crop
4. Moist agricultural field without crop
5. Eucalyptus Plantation
6. Broad leaved Forest

Also, field was visited to cross-check the classes on the ground. Though only north-eastern part (near Barakoli reservoir) was accessible, rest was quite remote and inaccessible. Still, it was manageable to recognize the classes on the image and give them appropriate label. The photographs taken during the field visit are given in the Figure 7.5.

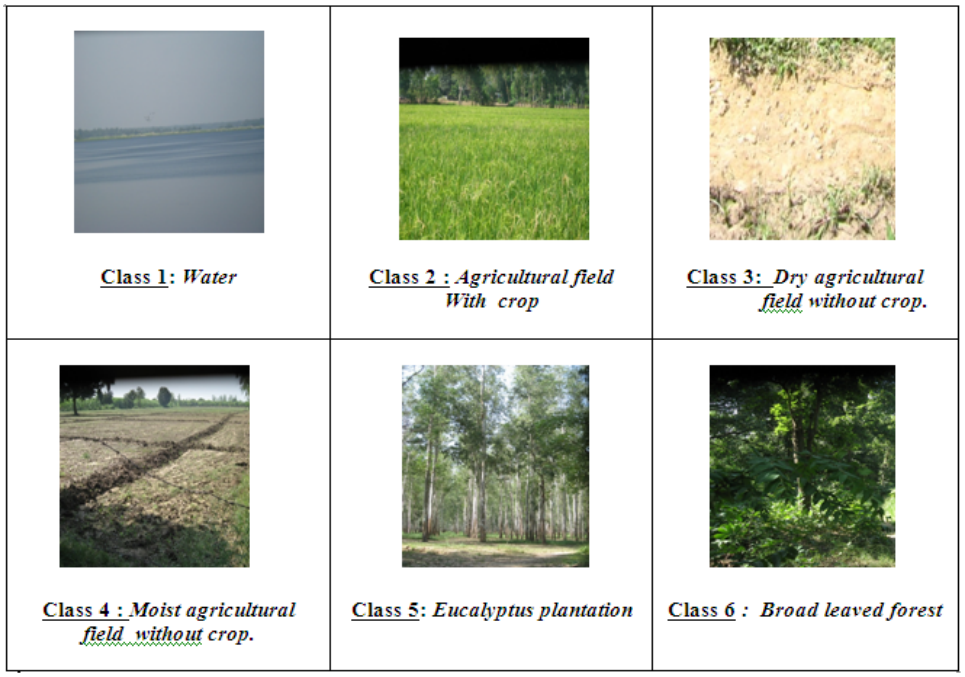


Figure 7.5: Photographs from the field showing different classes

Need of sub-pixel classification in the given study area Need of fuzzy based sub-pixel classification is quite evident on observing the images of AWIFS and LISS III (figure 7.6). Water body in the bottom right corner of the image is largely homogenous but except at the top which shows the presence of vegetation. This water class gradually changes into vegetation class. This was also evident during the field visit. The pixel size of AWIFS is more (60 m, after resampling) than that of LISS III (20m, after resampling), and thus, intergrade phenomena within pixel is more dominating in AWIFS than LISS III e.g. in case of water body and vegetation explained here.

The fields in this area are not very large, i.e. two or three pixels of AWIFS image (at maximum) are sufficient to cover an agricultural field. In the process, it also, includes the neighbouring class such as moist agricultural field without crop. Dry Agricultural fields without crop can be differentiated in LISS III images but in AWIFS, these classes also seem to merge into the neighbouring classes, especially at the boundary.

As observed from both the images and also, during the field visit the broad leaved forest and the eucalyptus plantation do not have a clear demarcation on the ground. Thus, as seen in Figure 7.6, the broad leaved forest (thick red patch at right side of the image) is not completely homogenous and dull red patches, within the broad leaved forest class, belonging to eucalyptus plantation are visible.

Thus, from the above discussion need of sub-pixel classification approach can be emphasised for areas having land cover classes, as taken up in this research.

Following subsections and further sections give the result of accuracy assessment. The reference images used here have been generated using two different classifiers, as discussed in section 5.3.4, chapter 5.

Since, there is no commercial tool available to assess the accuracy of sub-pixel classified output. Thus, JAVA based in house-tool developed by [51] has been used for assessing the accuracy (see Appendix B). It uses method of random sampling. The to-

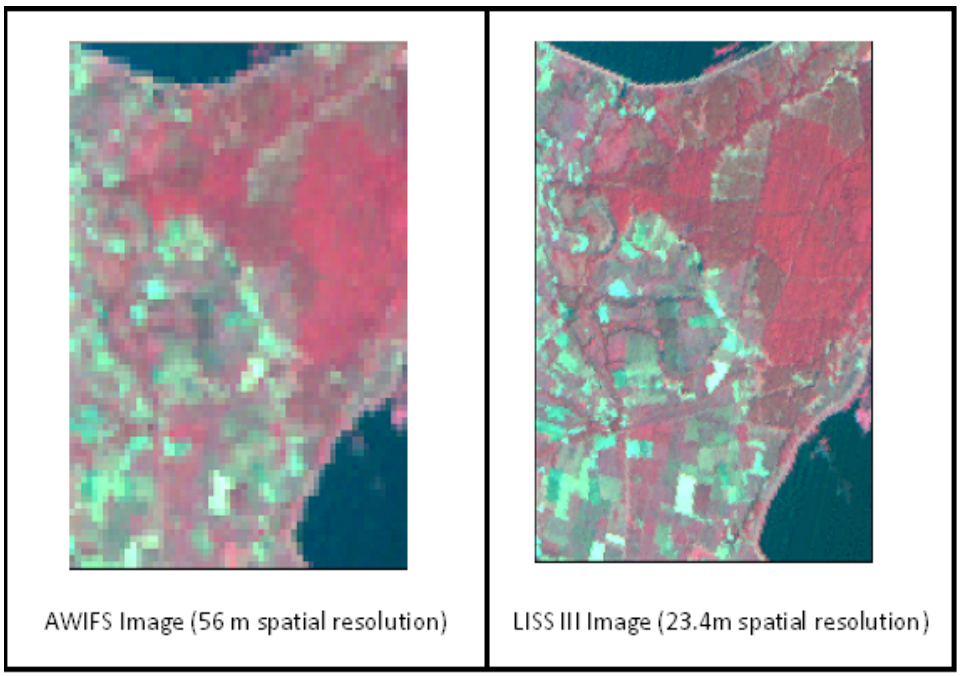


Figure 7.6: AWIFS and LISS III images of the study area

tal number of test pixels were taken based on Congalton's rule of 75 to 100 pixels, per class. Using the same software the training data was selected as 40 pixels per class in LISS IV, 20 pixels per class in LISS III and 8-10 pixels per class in AWIFS.

7.2.1 Performance of PCM-MRF I on coarser resolution dataset

This subsection discusses the results of classifying images from AWIFS. The output fraction images obtained on classification using PCM-MRF I are given in the Figure 7.7.

Results of FERM and MIN-MIN

Table 7.5 and 7.6 gives the result of the PCM-MRF I AWIFS classified data against the reference data set, LISS IV.

For the purpose of comparison, the AWIFS images were also classified using PCM. The results are given in Table 7.7 and 7.8. Figure 7.7 gives the results of fraction images generated using PCM.

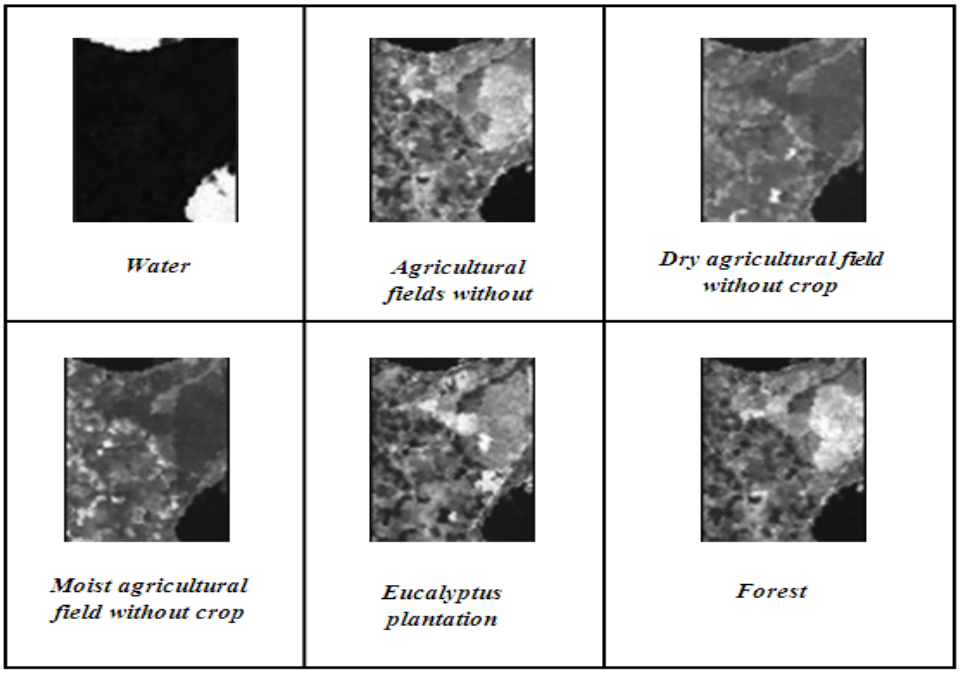


Figure 7.7: Resultant fraction images after classifying AWIFS with PCM-MRF I

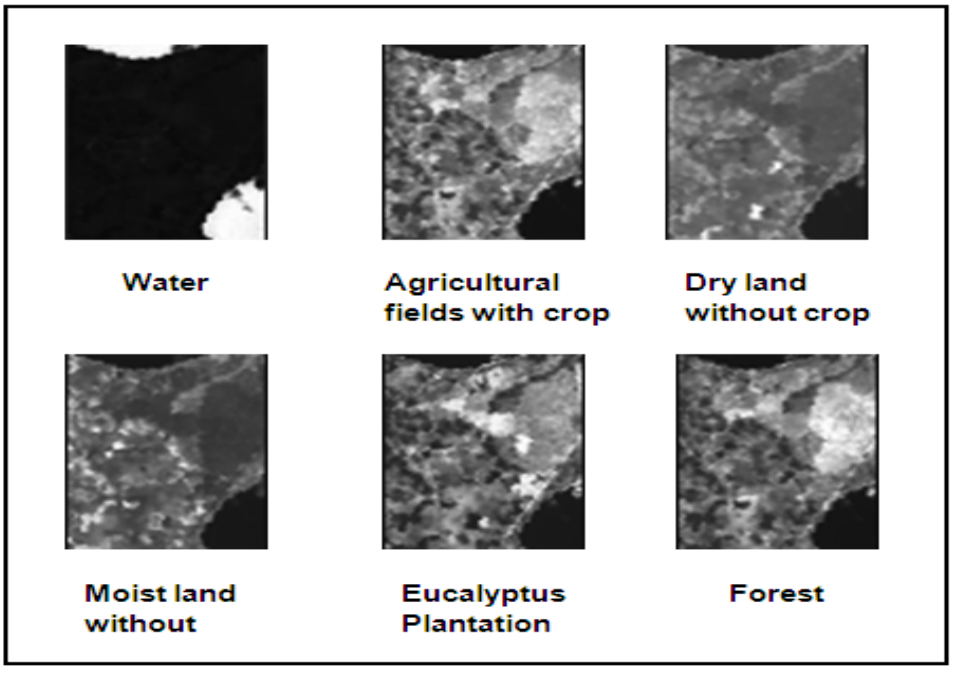


Figure 7.8: Resultant fraction images after classifying AWIFS with PCM

Table 7.5: Results of PCM-MRF I AWIFS classified data vs. PCM-MRF I classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	96.61	86.74
Producer's Accuracy	64.39	55.85
Agriculutral fields with crop		
User's Accuracy	95.04	92.02
Producer's Accuracy	85.92	86.11
Dry agricultural field without crop		
User's Accuracy	85.68	66.92
Producer's Accuracy	96.27	97.07
Moist agricultural field without crop		
User's Accuracy	89.65	73.05
Producer's Accuracy	92.93	91.01
Eucalyptus Plantation		
User's Accuracy	95.48	93.14
Producer's Accuracy	78.66	80.07
Forest		
User's Accuracy	95.75	93.45
Producer's Accuracy	85.31	84.21
Average User's Accuracy	93.03	
Average Producer's Accuracy	83.91	
Fuzzy Overall Accuracy	84.92	84.46
Fuzzy Kappa Coefficient		0.808

Table 7.6: Results of PCM-MRF I AWIFS classified data vs. SVM classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	68.3	72.31
Producer's Accuracy	88.81	74.45
Agriculutral fields with crop		
User's Accuracy	49.77	67.84
Producer's Accuracy	90.23	75.53
Dry agricultural field without crop		
User's Accuracy	40.46	52.29
Producer's Accuracy	98.83	95.55
Moist agricultural field without crop		
User's Accuracy	32.14	42.59
Producer's Accuracy	95.7	88.28
Eucalyptus Plantation		
User's Accuracy	87.58	95.88
Producer's Accuracy	70.94	42.96
Broad Leaved Forest		
User's Accuracy	22.02	37.5
Producer's Accuracy	99.85	99.29
Average User's Accuracy	50.04	
Average Producer's Accuracy	90.72	
Fuzzy Overall Accuracy	84.46	63.15
Fuzzy Kappa Coefficient		0.539

Table 7.7: Results of PCM classified AWIFS data vs. PCM classified reference data set

Accuracy Assessment Methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	96.884	88.06
Producer's Accuracy	59.95	52.12
Agriculutral field without crop		
User's Accuracy	95.05	92.07
Producer's Accuracy	84.13	84.42
Dry agricultural field without crop		
User's Accuracy	85.48	66.82
Producer's Accuracy	96.37	97.28
Moist agricultural field without crop		
User's Accuracy	89.39	72.90
Producer's Accuracy	92.78	90.97
Eucalyptus Plantation		
User's Accuracy	93.77	90.33
Producer's Accuracy	79.65	80.23
Broad Leaved Forest		
User's Accuracy	95.37	92.64
Producer's Accuracy	85.39	83.94
Average User's Accuracy	92.66	
Average Producer's Accuracy	83.06	
Fuzzy Overall Accuracy	84.35	83.14
Fuzzy Kappa Coefficient		0.795

Table 7.8: Results of PCM classified AWIFS data vs. SVM classified reference data set

	FERM (%)	**MIN-MIN (%)**
Water		
User's Accuracy	73.47	77.72
Producer's Accuracy	79.29	65.28
Agriculutral fields with crop		
User's Accuracy	48.45	66.33
Producer's Accuracy	90.73	77.34
Dry agricultural field without crop		
User's Accuracy	38.32	50.12
Producer's Accuracy	97.64	94.01
Moist agricultural field without crop		
User's Accuracy	30.42	40.22
Producer's Accuracy	93.82	85.09
Eucalyptus Plantation		
User's Accuracy	95.17	85.92
Producer's Accuracy	71.12	43.88
Broad Leaved Forest		
User's Accuracy	22.66	37.98
Producer's Accuracy	100	100
Average User's Accuracy	49.88	
Average Producer's Accuracy	88.77	
Fuzzy Overall Accuracy	83.24	62.71
Fuzzy Kappa Coefficient		.535

On comparing the fuzzy overall accuracy, in case of PCM-MRF I for AWIFS it was 84.46% with fuzzy kappa coefficient as .808 (from Table 7.5) where as in case of PCM it is 83.14% fuzzy overall accuracy and kappa coefficient of 0.795(Table 7.7). The increase was found to be 1.58% with increase in kappa coefficient to 1.61%. In case of SVM classified reference data set, the results of MIN-MIN does not seem to be consistent with that of FERM (Table 7.5 - 7.8). It may be because of change in classifier for reference data set. And as discussed in section 5.3.4, chapter 5, the use of different classifiers for reference and test data may cause inconsistency in the results. Also, on observing the results carefully, the users accuracy has drastically decreased to as low as 22.02%(table 7.6, class Broad Leaved Forest) where as same class had the user's accuracy of 95.75% when same classifier was used (table 7.5). Similar is the case for PCM (Table 7.7 and 7.8).

RMSE and r values

The RMSE and (r) values (Table 7.9 - 7.12) are given for the comparison.

Table 7.9: RMSE and r values for PCM-MRF I classified AWIFS data when reference data set was generated using PCM-MRF I

Class	RMSE	r
Water	0.56	0.83
Agricultural fields with crop	0.64	0.64
Dry agricultural field without crop	0.41	0.54
Moist agricultural field without crop	0.41	0.64
Eucalyptus plantation	0.73	0.59
Broad Leaved Forest	0.59	0.74
Global	0.23	0.76

Table 7.10: RMSE and r values for PCM MRF I classified AWIFS data when reference data set was generated using SVM

Class	RMSE	r
Water	0.57	0.83
Agricultural fields with crop	1.23	0.22
Dry agricultural field without crop	0.99	0.42
Moist agricultural field without crop	0.99	0.54
Eucalyptus plantation	1.07	0.38
Broad leaved Forest	1.41	0.58
Global	0.44	0.48

It is observed that in case of PCM a high RMSE of more than one is seen in case of agricultural fields with crops, Eucalyptus plantation and Broad Leaved Forest (table 7.11). Whereas, in PCM-MRF I, all RMSE for all three above mentioned classes have been considerably reduced (table 7.9).

In case of PCM-MRF I lower RMSE (0.23) and r(0.76) was found as compared to PCM (RMSE = 0.40, r = 0.69). This is in consistent to the results obtained from FERM and MIN-MIN.

Table 7.11: RMSE and r values for PCM classified AWIFS data when reference data set was generated using PCM

Class	**RMSE**	**r**
Water	0.71	0.82
Agricultural fields with crop	1.10	0.57
Dry agricultural field without crop	0.87	0.48
Moist agricultural field without crop	0.92	0.59
Eucalyptus plantation	1.13	0.55
Broad Leaved Forest	1.07	0.70
Global	0.40	0.69

Table 7.12: RMSE and r values for PCM classified AWIFS data when reference data set was generated using SVM

Class	RMSE	r
Water	0.57	0.83
Agricultural fields with crop	1.23	0.22
Dry agricultural field without crop	0.99	0.42
Moist agricultural field without crop	0.99	0.58
Eucalyptus plantation	1.41	0.38
Broad Leaved Forest	0.59	0.74
Global	0.44	0.48

Entropy

In this research, entropy was calculated for the experimental purpose to see the absolute uncertainty of the classified output. Lower entropy value implies lesser uncertainty and better classifier [47]. From these experiment entropy for PCM-MRF I found to be 2.3057, whereas in case of PCM it was 2.3068.

7.2.2 Performance of PCM-MRF I on moderate resolution images

Figure 7.9 shows the classified output of LISS III images using PCM-MRF I.

Results of FERM and MIN-MIN

Table 7.13 and 7.14 gives the result of accuracy assessment of PCM-MRF I classified output of PCM-MRF I with reference data generated from LISS IV.

These results have been compared with the results of PCM classified LISS III images, given in table 7.15 and 7.16. Also, figure 7.10 gives the PCM classified LISS III output.

Comparing the results of PCM-MRF I (Table 7.13,7.14) of LISS III with results of PCM (table 7.15, 7.16), it is found that the accuracy of PCM-MRF I classified LISS III images is slightly less than that of PCM. The results using MIN-MIN operator give fuzzy kappa coefficient as 0.889 for PCM-MRF I with fuzzy overall accuracy of 90.97% whereas for PCM the fuzzy kappa coefficient is 0.891 and fuzzy overall accuracy is 91.04%. This trend is visible in all the other results as well in all the four tables (Table

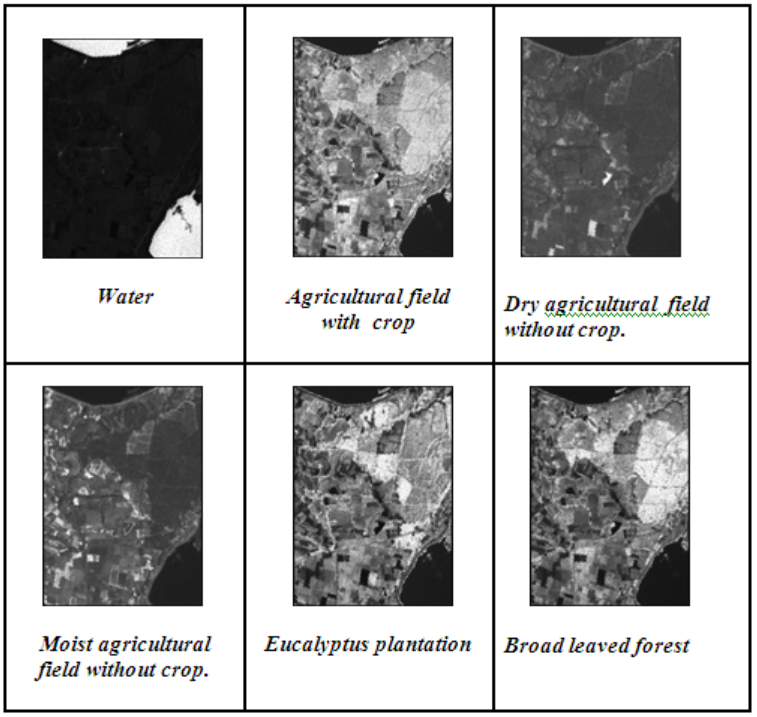


Figure 7.9: Resultant fraction images after classifying LISS III image with PCM-MRF I

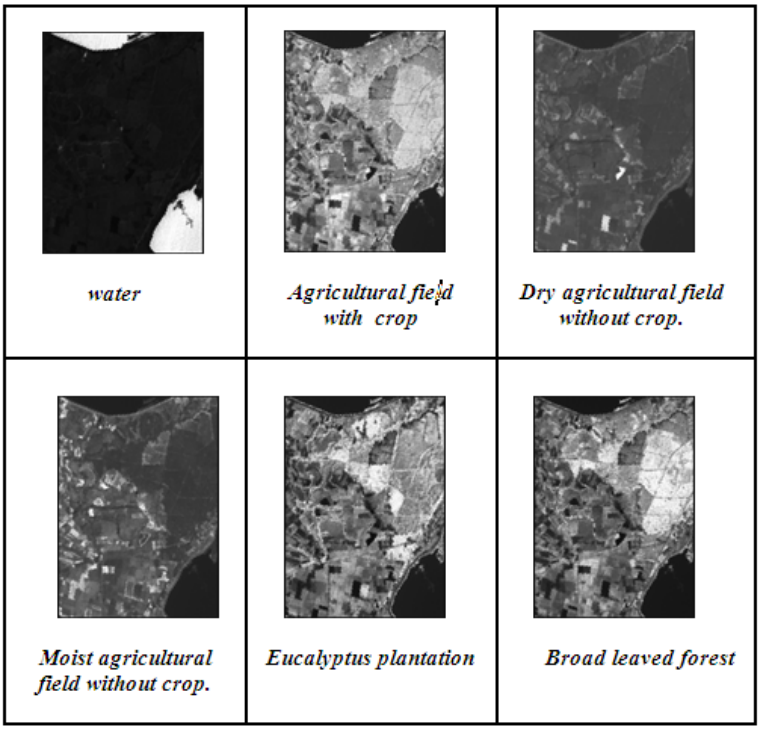


Figure 7.10: Resultant fraction images on classifying LISS III image using PCM

Table 7.13: Results of PCM-MRF I classified LISS III data vs. PCM-MRF I classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	97.91	95.53
Producer's Accuracy	74.02	72.51
Agriculutral fields with crop		
User's Accuracy	88.83	88.83
Producer's Accuracy	92.63	96.53
Dry agricultural field without crop		
User's Accuracy	99.14	98.17
Producer's Accuracy	87.01	86.55
Moist agricultural field without crop		
User's Accuracy	94.56	84.98
Producer's Accuracy	92.63	92.40
Eucalyptus Plantation		
User's Accuracy	93.23	89.69
Producer's Accuracy	87.46	92.88
Broad Leaved Forest		
User's Accuracy	96.08	93.06
Producer's Accuracy	89.49	93.93
Average User's Accuracy	95.60	
Average Producer's Accuracy	87.21	
Fuzzy Overall Accuracy	88.48	90.97
Fuzzy Kappa Coefficient		0.889

7.13 to 7.16). This slight difference may be attributed to random sampling. Ignoring this small difference, it is observed that results have not improved by including spatial context with PCM using standard regularizer. As discussed in section 4.3.4, chapter 4, standard regularizer performs over smoothing and hence, inspite of adding contextual information in PCM-MRF I, the results did not improved.

Table 7.14: Results of PCM-MRF I classified LISSIII data vs. SVM classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	70.91	76.81
Producer's Accuracy	78.99	64.66
Agriculutral fields with crop		
User's Accuracy	42.29	61.87
Producer's Accuracy	89.97	76.55
Dry agricultural field without crop		
User's Accuracy	46.26	58.35
Producer's Accuracy	97.64	94.01
Moist agricultural field without crop		
User's Accuracy	31.63	43.11
Producer's Accuracy	93.13	85.17
Eucalyptus Plantation		
User's Accuracy	73.57	47.19
Producer's Accuracy	81.53	93.07
Broad Leaved Forest		
User's Accuracy	21.85	38.81
Producer's Accuracy	99.7	98.6
Average User's Accuracy	49.68	
Average Producer's Accuracy	88.25	
Fuzzy Overall Accuracy	83.96	63.98
Fuzzy Kappa Coefficient		0.548

Table 7.15: Results of PCM classified LISS III data vs. PCM classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	98.26	94.31
Producer's Accuracy	74.27	73.49
Agriculutral fields with crop		
User's Accuracy	92.34	88.04
Producer's Accuracy	92.97	97.28
Dry agricultural field without crop		
User's Accuracy	86.69	98.38
Producer's Accuracy	99.20	85.70
Moist agricultural field without crop		
User's Accuracy	95.49	87.56
Producer's Accuracy	92.28	90.7
Eucalyptus Plantation		
User's Accuracy	93.42	90.04
Producer's Accuracy	88.06	93.05
Broad Leaved Forest		
User's Accuracy	95.32	91.92
Producer's Accuracy	90.19	95.03
Average User's Accuracy	95.67	
Average Producer's Accuracy	87.41	
Fuzzy Overall Accuracy	88.58	91.04
Fuzzy Kappa Coefficient		0.891

Table 7.16: Results of PCM classified LISS III data vs. SVM classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	65.72	71.49
Producer's Accuracy	80.36	65.57
Agriculutral fields with crop		
User's Accuracy	43.93	62.34
Producer's Accuracy	90.42	78.04
Dry agricultural field without crop		
User's Accuracy	48.89	60.71
Producer's Accuracy	92.83	85.13
Moist agricultural field without crop		
User's Accuracy	33.64	45.25
Producer's Accuracy	92.15	84.38
Eucalyptus Plantation		
User's Accuracy	83.16	93.39
Producer's Accuracy	73.95	46.89
Broad Leaved Forest		
User's Accuracy	19.49	35.35
Producer's Accuracy	99.74	98,81
Average User's Accuracy	49.14	
Average Producer's Accuracy	88.41	
Fuzzy Overall Accuracy	84	64.14
Fuzzy Kappa Coefficient		0.553

RMSE and r values

The Table 7.17 and 7.18 gives the RMSE and r values of PCM-MRF I LISS III classified output with respect to reference data sets. For comparison RMSE and r value of PCM has also been given with the same reference data set (Table 7.19 and 7.20).

Table 7.17: RMSE and r values for PCM-MRF I classified LISS III data w.r.t PCM-MRF I classified reference data set

Class	**RMSE**	**r**
Water	0.27	0.47
Agricultural fields with crop	0.29	0.31
Dry agricultural field without crop	0.15	0.19
Moist agricultural field without crop	0.19	0.28
Eucalyptus plantation	0.30	0.34
Broad Leaved Forest	0.28	0.41
Global	0.10	0.55

Table 7.18: RMSE and r values for PCM-MRF I classified LISS III data w.r.t reference data from SVM

Class	**RMSE**	**r**
Water	0.31	0.48
Agricultural fields with crop	0.50	0.06
Dry agricultural field without crop	0.28	0.13
Moist agricultural field without crop	0.35	0.22
Eucalyptus plantation	0.43	0.20
Broad Leaved Forest	0.55	0.26
Global	0.17	0.31

Table 7.19: RMSE and r values for PCM classified LISS III data w.r.t PCM classified reference data

Class	**RMSE**	**r**
Water	0.28	0.47
Agricultural fields with crop	0.28	0.31
Dry agricultural field without crop	0.15	0.19
Moist agricultural field without crop	0.19	0.28
Eucalyptus plantation	0.30	0.34
Broad Leaved Forest	0.28	0.41
Global	0.10	0.55

Table 7.20: RMSE and r values for PCM classified LISS III data w.r.t SVM classified reference data

Class	**RMSE**	**r**
Water	0.31	0.48
Agricultural fields with crop	0.50	0.07
Dry agricultural field without crop	0.27	0.13
Moist agricultural field without crop	0.35	0.23
Eucalyptus plantation	0.43	0.21
Broad Leaved Forest	0.55	0.26
Global	0.17	0.55

The result of RMSE and r of PCM-MRF I, 0.10 and 0.55 (from table 7.17) was same for PCM (table 7.19). Similarly, results from table 7.18 and table 7.20 (reference data was SVM) follow a same trend, qualitatively. Thus, it substantiates the results obtained from FERM and MIN-MIN operators.

entropy

The resultant entropy for PCM was found to 2.3224 of PCM-MRF I with entropy being 2.3220.

From the results observed, it can be said there is no significant improvement on using standard regularizer for including contextual information in PCM. It can be attributed to the functioning of standard regularizer, which performs over smoothening as discussed in section 4.3.1, chapter 3.

Also, results of LISS III are different from that of AWIFS. In other words, PCM-MRF I has shown improvement in case of AWIFS but in case of LISS III there is not significant difference between PCM and PCM-MRF I. The effect of over smoothening by standard regularization method is more in LISS III, as compared to AWIFS.

7.3 Performance of DA based PCM-MRF : PCM-MRF II

The aim behind using DA model for including prior information was to preserve the edges.

7.3.1 Performance of PCM-MRF II on coarse resolution images

Method to verify edge preservation

To verify the preservation of edges, a homogenous area of a class was selected by verifying that it should have high mean and low variance.

After identifying homogenous area, two set of pixels were selected with each set corresponding to either side of the edge. Mean and variance of these two were calculated using ENVI 4.3. Comparative (PCM-MRF II vs. PCM-MRF I) difference of the mean between these two set of pixel should be high when edge is preserved. Comparative (PCM-MRF II vs. PCM-MRF I) variance within each pixel set should be low, as it describes low variability within the set.

Table 7.21 gives the result of mean and variance calculated of various classes, as stated above.

Results in Table 7.21, shows that edge are better preserved in case of PCM-MRF II as compared to PCM-MRF I. As per [32], the error usually occurs around the edges, and

Table 7.21: Verifying preservation of edges using Mean and variance for AWIFS image

Class	**PCM-MRF I**		**PCM-MRF II**	
	Diff. in mean	**Variance**	**Diff. in mean**	**Variance**
Water	171.80	1604.66, 388.54	222.45	1001.82, 375.8
Agriculutral fields with crop	98.63	117.29, 292.41	107.23	83.33, 228.01
Dry agricultural without crop	105.18	256.32, 75.34	110.91	272.25, 56.25

thus, on preserving the edges, an increase in overall accuracy should also be observed. This can be verified from the results of accuracy assessment of PCM-MRF II given in Table 7.22.

Table 7.22 and 7.23 gives the accuracy assessment of PCM-MRF II classified with reference data set generated using SVM and PCM-MRF II, respectively. It is seen the fuzzy kappa coefficient is greater in case of PCM-MRF II ($\kappa = 0.819$) as compared to that of PCM-MRF I ($\kappa = 0.808$) (from table 7.18 and 7.1, respectively), where the increment is 1.4%. The resultant increase in value of kappa coefficient is 3% as compared to that of PCM (0.795). Also, the overall accuracy from MIN-MIN operator is 85% where as for PCM-MRF I it is 84.46% and PCM is 83.14 %.

Thus, it is observed that the accuracy of PCM-MRF II is higher than that of PCM and also, has preserved the edges.

7.3.2 RMSE and r

Table 7.24 gives the RMSE and r values for PCM-MRF II. RMSE of PCM-MRF II was 0.21 which is the least in all the three case 0.23 PCM-MRF I and 0.40 for PCM. Similarly, results of r was highest in case of PCM-MRF II.

Entropy value

The Average entropy was found to be 2.3050 which is less than PCM-MRF I(2.3057) and PCM(2.3068). The entropy was also found to be less in case of PCM-MRF II as compared to other two classifiers.

Table 7.22: Results of PCM-MRF II AWIFS classified data vs. PCM-MRF II classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	97.35	88.67
Producer's Accuracy	64.96	59.77
Agriculutral fields with crop		
User's Accuracy	95.14	91.57
Producer's Accuracy	86.37	86.93
Dry agricultural field without crop		
User's Accuracy	86.57	68.13
Producer's Accuracy	96.77	97.84
Moist agricultural field without crop		
User's Accuracy	90.44	75.24
Producer's Accuracy	92.14	90.91
Eucalyptus Plantation		
User's Accuracy	96.7	93.55
Producer's Accuracy	79.31	80.99
Broad Leaved Forest		
User's Accuracy	95.17	92.89
Producer's Accuracy	84.84	83.23
Average User's Accuracy	94.19	
Average Producer's Accuracy	84.98	
Fuzzy Overall Accuracy	85.54	85
Fuzzy Kappa Coefficient		0.81

Table 7.23: Results of PCM-MRF II classified AWIFS data vs. SVM classified reference data set

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	75.69	68.75
Producer's Accuracy	84.32	80.08
Agriculutral fields with crop		
User's Accuracy	49.36	69.1
Producer's Accuracy	90.03	75.6
Dry agricultural field without crop		
User's Accuracy	40.35	53.18
Producer's Accuracy	98.21	94.97
Moist agricultural field without crop		
User's Accuracy	31.93	42.88
Producer's Accuracy	94.2	84.74
Eucalyptus Plantation		
User's Accuracy	72.98	45.38
Producer's Accuracy	85.43	95.74
Broad Leaved Forest		
User's Accuracy	23.53	30.55
Producer's Accuracy	99.8	99.03
Average User's Accuracy	51.05	
Average Producer's Accuracy	89.92	
Fuzzy Overall Accuracy	85.02	64.66
Fuzzy Kappa Coefficient		.562

Table 7.24: RMSE and r values for PCM-MRF II classified AWIFS image w.r.t PCM-MRF II classified reference data

Class	**RMSE**	**r**
Water	0.46	0.83
Agricultural fields with crop	0.46	0.71
Dry agricultural field without crop	0.33	0.59
Moist agricultural field without crop	0.36	0.74
Eucalyptus plantation	0.63	0.65
Broad Leaved Forest	0.42	0.78
Global	0.21	0.79

7.3.3 Performance of PCM-MRF II on moderate resolution images

Mean and variance was also calculated in case of LISS III images for verifying edge preservation. The results are given in table 7.25.

Table 7.25: Verifying preservation of edges using Mean and variance for LISS III image

Class	**PCM-MRF I**		**PCM-MRF II**	
	Diff. in mean	**Variance**	**Diff. in mean**	**Variance**
Water	152.05	129.96, 214.92	162.7	97.81, 177.69
Agriculutral field with crop	171.84	142.25, 179.56	177.86	112.04, 158.76
Dry agricultural field without crop	139.7	256.32, 75.34	144.8	175.29, 11.42

Also, the overall accuracy of PCM-MRF II was calculated to assess the performance of PCM-MRF II. The results are given in Table 7.26 and 7.27.

Table 7.26: Results of PCM-MRF II classified LISS III data vs. PCM-MRF II classified reference data

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	98.93	95.26
Producer's Accuracy	77.02	74.57
Agriculutral fields with crop		
User's Accuracy	93.23	88.73
Producer's Accuracy	93.78	97.01
Dry agricultural field without crop		
User's Accuracy	88.14	98.66
Producer's Accuracy	98.64	86.73
Moist agricultural field without crop		
User's Accuracy	96.11	87.85
Producer's Accuracy	93.05	91.72
Eucalyptus Plantation		
User's Accuracy	93.81	90.77
Producer's Accuracy	87.04	93.03
Broad Leaved Forest		
User's Accuracy	95.17	92.44
Producer's Accuracy	91.12	96.03
Average User's Accuracy	96.71	
Average Producer's Accuracy	88.11	
Fuzzy Overall Accuracy	89.45	92.01
Fuzzy Kappa Coefficient		0.904

Table 7.27: Results of PCM-MRF II classified LISS III data vs. SVM classified reference data

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	81.99	65.66
Producer's Accuracy	70.91	76.81
Agriculutral fields with crop		
User's Accuracy	43.46	62.58
Producer's Accuracy	89.25	79.08
Dry agricultural field without crop		
User's Accuracy	47.14	60.69
Producer's Accuracy	93.81	86.9
Moist agricultural field without crop		
User's Accuracy	31.44	45.43
Producer's Accuracy	94.64	87.01
Eucalyptus Plantation		
User's Accuracy	76.57	50.71
Producer's Accuracy	80.42	92.29
Broad Leaved Forest		
User's Accuracy	22.66	37.98
Producer's Accuracy	100	100
Average User's Accuracy	50.69	
Average Producer's Accuracy	89.04	
Fuzzy Overall Accuracy	84.98	66.27
Fuzzy Kappa Coefficient		.577

As observed from table 7.20, the edges are preserved in case of PCM-MRF II as well when compared with PCM-MRF I. The overall accuracy of PCM-MRF II(92.01%) is also higher than PCM-MRF I(91.04%) and PCM (90.97%). Also, RMSE and r values are provided in table 7.23 and 7.24.

RMSE and r values

Entropy

The entropy of PCM-MRF II was 2.3210, which was minimum as compared 2.3220 of PCM-MRF I and 2.3224 of PCM.

7.4 What if there are untrained classes?

Foody [5], showed that supervised PCM works better in case of untrained classes than supervised FCM. In this research work, the performance of supervised PCM-MRF was assessed in case of untrained classes and compared with that of PCM. To obtain the untrained class, the mean value of one of the classes was not provided to the classifier and the classification was done for remaining five classes as done by [5]. RMSE and r values were calculated to check the performance of PCM-MRF in case of untrained classes.

Table 7.28: RMSE and r values for PCM-MRF II classified LISS III data when reference data set was generated using PCM-MRF II

Class	**RMSE**	**r**
Water	0.25	0.52
Agricultural fields with crop	0.24	0.38
Dry agricultural field without crop	0.11	0.25
Moist agricultural field without crop	0.13	0.34
Eucalyptus plantation	0.24	0.38
Broad Leaved Forest	0.23	0.44
Global	0.08	0.61

7.4.1 Discussion on results

The results of assessment of AWIFS and LISS III images with untrained classes are shown in Table 7.29 - Table 7.32, respectively. The results consists of one case from AWIFS when class 5 i.e. eucalyptus plantation was missing. For LISS III, two cases were taken as when class 4 i.e. moist agricultural land without crop was missing and second was class 5 i.e. eucalyptus plantation was left untrained.

Table 7.29: RMSE and r values of AWIFS, when class 5 (Eucalyptus plantation) was missing

	PCM		PCM-MRF I		PCM-MRF II	
	RMSE	r	RMSE	r	RMSE	r
Water	0.3077	0.4627	0.3072	0.463	0.3073	0.4628
Agricultural field with crop	0.4431	0.0834	0.4433	0.0833	0.4432	0.0836
Dry agricultural field without crop	0.3536	0.2382	0.3538	0.2375	0.3536	0.2381
Moist agricultural field without crop	0.4571	0.4073	0.4572	0.4068	0.4572	0.4066
Broad leaved forest	0.29	0.135	0.2900	0.1344	0.29	0.135
Global	0.1684	0.41	0.1684	0.4098	0.1684	0.4099

The resultant values of RMSE and r, in case of untrained classes, observed in table 7.29 to 7.31, are either same or similar(when observed uptil 4th decimal point). PCM performs better than FCM in case of untrained classes because it is independent of the number of classes present in the image, unlike FCM [6, 5]. The results obtained here show that using MRF doesn't degrade the performance of the PCM in case of untrained classifier, as the results obtained are same.

7.5 Summary

The results obtained from the accuracy assessment of the classified output implies that including contextual information improves the result of PCM based classifier especially, in case of coarser resolution data such as AWIFS. But accuracy results differ for the two different prior models used and on preserving the edges the accuracy assessment

results of the classifier improves as seen in case of PCM-MRF II for both AWIFS and LISS III images.

But as discussed in section 5.3.1, the resampling induces errors in the images which are inevitable. Thus, error in classified output cannot be completely attributed to the performance of the classifier. Also, due to the limitation of the accuracy assessment tool used, that it used random sampling, the performance of PCM-MRF II at the edges in particular, cannot be measured fully.

The next chapter i.e. chapter 8 gives the conclusion based on the results and discussion given in this chapter.

Table 7.30: RMSE and r values of LISS III, when class 4 (moist agricultural field without crop) was left untrained

	PCM		PCM-MRF I		PCM-MRF II	
	RMSE	r	RMSE	r	RMSE	r
Water	0.32	0.4793	0.3193	0.4798	0.3189	0.4794
Agricultural field with crop	0.5277	0.046	0.528	0.0458	0.5283	0.046
Dry agricultural field without crop	0.613	0.2574	0.6132	0.256	0.6128	0.2565
Eucalyptus plantation	0.4572	0.1579	0.4577	0.1569	0.4579	0.1567
Broad leaved forest	0.4572	0.1579	0.4577	0.1569	0.4579	0.1567
Global	0.2077	0.3204	0.2078	0.32	0.2077	0.32

Table 7.31: RMSE and r values of LISS III, when class 5 (Eucalyptus plantation) was missing

	PCM		PCM-MRF I		PCM-MRF II	
	RMSE	r	RMSE	r	RMSE	r
Water	0.3077	0.4627	0.3072	0.463	0.3073	0.4628
Agricultural field with crop	0.4431	0.0834	0.4433	0.0833	0.4432	0.0836
Dry agricultural field without crop	0.3536	0.2382	0.3538	0.2375	0.3536	0.2381
Moist agricultural field without crop	0.4571	0.4073	0.4572	0.4068	0.4572	0.4066
Broad leaved forest	0.29	0.135	0.2900	0.1344	0.29	0.135
Global	0.1684	0.41	0.1684	0.4098	0.1684	0.4099

Chapter 8

Conclusion and Recommendations

The chapter has been bifurcated into conclusion (section 8.1) and recommendations (section 8.2). Section 8.1 has been further divided into :

1. Answers to the research questions posed in the beginning of this research. (section 8.1.1)
2. Points of conclusion (section 8.1.2)

Section 8.2, discusses the recommendations that can and must be carried forward as further research.

8.1 Conclusion

The main objective of this research was to develop a sub-pixel classifier for classifying moderate and coarse spatial resolution multi-spectral data set using PCM and MRF. Also, to gauge the performance of the classifier that uses spatial contextual information against PCM based classifier, which uses only spectral information and no spatial context is included.

Another objective was to use edge-preserving priors i.e DA models, to preserve the edges. The efficacy of the same has been discussed in section 7.3, chapter 7 in terms of the results.

A method was developed which incorporated spatial context using both standard regularization model and DA model and performed sub-pixel classification in supervised mode. And it was also observed that PCM-MRF with DA model preserved the edges.

8.1.1 Answers to the research questions.

Research Q1. **How to modify the PCM objective function so as to include the spatial contextual information using MRF?**

Answer : The answer to this research question is the objective functions developed and discussed in section 6.1, chapter 6. In section 6.1.1, equation 6.4 gives the PCM-MRF which includes standard regularization model and in section 6.1.2, equation 6.5, includes DA model for edge preservation. These equations were applied and tested on the data set used in this research. Also, the resultant images as well as their accuracy assessment results were discussed in section 7.

From the results and discussion (chapter 7) it can be concluded that :

- The proposed method is one of the means by which contextual information can be added with PCM for performing sub-pixel classification.
- Adding the spatial contextual information (using MRF) in PCM does improves the accuracy of the classifier $(by\ 1.5\% - 3\%)$.
- DA priors in PCM-MRF preserves the edges which improves the overall accuracy of the classifier which is consistent with the literature surveyed [32, 17].

Research Q2. **Which smoothness prior model will be suitable for PCM-MRF?**

Answer: In this research work two smoothness models were used, as discussed in section 4.2, chapter 4.

1. Standard Regularizer (equation 4.7)
2. Discontinuity Adaptive Model (equation 4.14))

As discussed in section 4.3, chapter 4, standard regularizers assume constant interaction among the pixels in image space, which may lead to over-smoothening. On observing the results of mean and variance (section 7.3) calculated at the edges and the accuracy assessment (section 7.2 and 7.3), it can be concluded that standard regularizers (PCM-MRF I) are not able to preserve the edges as efficiently as DA models used in PCM-MRF II.

As well, on preserving the edges, accuracy of the classifier increases by 2%.

Research Q3. **Which sampler will be suitable for PCM-MRF?**

Answer: To draw random samples from a distribution, Markov Chain Monte Carlo methods (MCMC) are employed. The Metropolis-Hastings Sampler and the Gibbs Sampler are two most fundamental MCMC techniques. In both the cases, at a time 't' the next state i.e $f(t+1)$, is generated as per Markov transition probability given by $P(f(t+1)|f(t))$.

The difference between the Metropolis-Hastings and Gibbs Sampler lies in the method adopted to generate $f(t+1)$. In Metropolis-Hastings sampler (in general), and Metropolis sampler (in particular), the next configuration is a random choice from $N(f)$, $N(f)$ being the discreet set of sample space. It is not suitable to be used when the sample space is infinite and the data is of high dimensionality. Also, seen in the work of Dutta [13].

Where as, in case of Gibbs Sampler it is chosen from the conditional distribution (when upto pairwise cliques are used) given by equation 8.1 [21, 52] :-

$$P(f_i|f_{\mathcal{N}_i}) = \frac{e^{-[V_1(f_i)+\sum_{i'\epsilon\mathcal{N}_i} V_2(f_i,f_i')]/T}}{\sum_{f_i\epsilon\mathcal{L}} e^{-[V_1(f_i)+\sum_{i'\epsilon\mathcal{N}_i} V_2(f_i,f_i')]/T}} \tag{8.1}$$

Hence, Gibbs Sampler generate the values from a pdf obtained from equation 8.1 and therefore, can work when sample space have infinite size and with high dimensionality data as in case of sub-pixel classifier where the membership value lies between 0 to 1. Also, seen in previous work by [13].

Hence, it is concluded that for a sub-pixel classifier Gibbs Sampler is the suitable sample over Metropolis sampler.

Research Q4. **Which classifier to use for generation of reference data set from LISS IV images?**

Answer: As stated in section 5.3.4, chapter 5, the generation of reference data is more preferable using sub-pixel classifier instead of using the method of multiple resolution, when the classifier to be tested generates the fuzzy membership values.

Also, there was an issue of selecting a suitable classifier for generating reference data set. In section 5.3.4 it was discussed, that using the different classifier may induce error due to difference in spatial resolution and lack of knowledge on the performance of the classifier for reference data set with respect to the classifier used for the test images. Also, as observed in the results of chapter 7 (Table 7.2, 7.4, 7.10, 7.12) the user and producer accuracy were quite inconsistent with respect to results obtained from other classifier for reference data set (Table 7.1, 7.3, 7.9, 7.11). As well there was inconsistency between the FERM results and MIN-MIN operator results when different was used for generating reference data set.

Due to these inconsistency observed, the results obtained when reference data set was generated using different classifier cannot be used. Hence, it is appropriate to use same classifier for generating reference data set as well as test data set.

Research Q5. **Which accuracy assessment technique is suitable to assess the accuracy of PCM-MRF sub-pixel classifier?**

Answer: Out of the available techniques discussed in section 6.4.1, chapter 6 and substantiated by the results obtained (see Appendix B), only FERM and MIN-MIN were found suitable for assessing the accuracy of PCM based classifier such as PCM-MRF.

Thus, it can be concluded that, methods which are not constrained by the hyperline constraint (equation 1.1), are suitable for assessing the accuracy of PCM based sub-pixel classifiers. And other classifiers such as SCM are not suitable as PCM do not follow the hyperline constraint.

Research Q6. Which accuracy assessment technique will be suitable to assess the accuracy of PCM-MRF at the edges?

Answer : As seen in section 7.3.1, chapter 7, to see whether the edge was preserved i.e. the edge information was not degraded, difference in mean of the pixels along the edges was measured. To further confirm the results, it is important to consider the overall performance (in terms of overall accuracy,kappa coefficient) of the classifiers as well [32].

Research Q7. How does PCM-MRF perfoms in case of noise/ outliers/ untrained classes?

Answer : case 1 : Untrained classes PCM-MRF I and PCM-MRF II was tested for untrained classes. RMSE and r values were calculated for PCM,PCM-MRF I and PCM-MRF II and from the results obtained (section 7.4), it can be seen that these values were same for PCM-MRF I, PCM-MRF II and PCM . It can be concluded that performance of PCM in presence of untrained classes is not reduced when contextual information is used.

case 2 : Presence of noise and/or outliers To assess the performance of PCM-MRF in case of noise and outliers, unsupervised classification method is required. In unsupervised classification, the mean values/training set is calculated using the iterative methods, with minimum intervention from the analyst. Noise and outliers causes problem in calculating the correct cluster mean values [6, 11].

But in this research work, the supervised mode is being used i.e. the analysts provides the training data. Thus, performance of PCM-MRF classifier cannot be measured in case of noise and outliers, in this research work.

Thus, it can be concluded that performance of classifier in presence of noise and outliers can be efficiently tested in unsupervised mode.

8.1.2 Further Conclusions and potential user/users

1. It has been found that k is different for AWIFS and LISS III. It was observed that it increase with decrease in spatial resolution of the image and with the increase in the number of pixels in the image. It was also seen that, by decreasing the size of the image and number of classes in the image (keeping the spatial resolution constant) the update rate decreased. Hence, it can be concluded that update rate (for image classification) may depend on factors such as spatial resolution, image size and number of classes.

   Note: It requires rigrous testing and further studies to be done to assess the effect of these parameteres on update rate.

2. The value of λ for LISS III was $\lambda = 0.4$ and for AWIFS was $\lambda = 0.8$. It helps in concluding that role of spatial context is more in case of coarser resolution and as the resolution of the image becomes finer the role of spatial context decreases.

3. Referring back to the results of PCM-MRF I and PCM-MRF II in case of AWIFS and LISS III, it was found that PCM-MRF I and II shows better performance as compared to that of PCM for coarser resolution data. But in case of moderate resolution i.e. LISS III the accuracy of PCM-MRF I did not showed improvement, that can be attributed to the increase in the discontinuities, as the image is more fine as compared to AWIFS. Where as, when DA priors were used for LISS III classification, there was rise in the accuracy.

   It shows, as the spatial resolution becomes finer the need of preserving the edges becomes essential.

4. As per the entropy calculated, the uncertainty of the results decreased in case of PCM-MRF I and further, in PCM-MRF II. These uncertainty results were independent of the reference data set and hence, can be said are absolute results. Further, rigorous experiments are recommended to be performed to see the amount of difference obtained in entropy values.

The conclusion is incomplete without defining the potential users of the sub-pixel classification method.

PCM-MRF gives better accuracy than popular fuzzy based classifiers, by including spatial contextual information and overcoming hyperline constraint. Thus, it can be used to generate spatially and spectrally consistent thematic maps. These thematic maps are important for resource management, which fulfils the purpose of launching satellites such as IRS-P6. A user if interested in single class, that can also be extracted from PCM based methods. Using PCM-MRF II a single class of interest can be extracted while preserving the edges.

8.2 Recommendations

The thesis is incomplete without further recommendations. It is my observation that this research is just the tip of the ice-berg and lot more can and should be done to completely exploit the potential of PCM-MRF . Figure 8.2 broadly defines the steps

followed in this thesis. The points of recommendation have been given in a sequence, corresponding to the steps in figure 8.2.

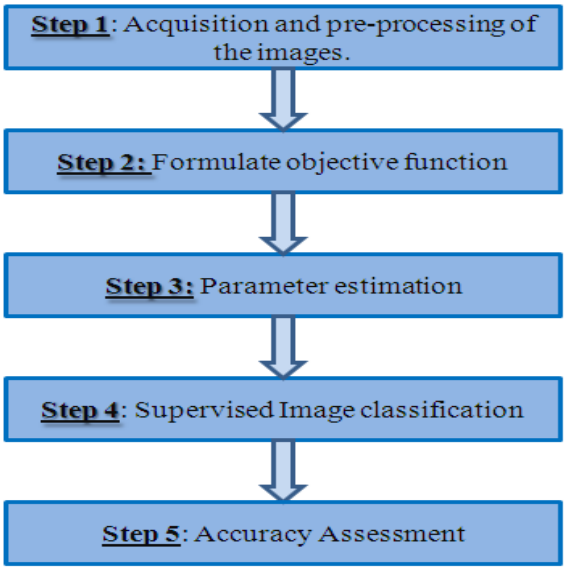


Figure 8.1: Flow chart showing the steps of methodology broadly

1. Beginning from step 2.
    - Instead of using objective function itself, the formula for calculating membership value from the objective function can be derived. The classification results obtained using the membership values should be compared with the present method of using objective function. And the assess the difference, if any.
    - In the PCM-MRF II, other three DA models by Li, can also be included.
2. Step 3 : parameter estimation.
    - More rigrous estimation of parameters with various methods (such as cross validation)is needed. Currently, there is no method which can be termed as standard, for estimating these parameters. A comparative study can be done for all the methods available and possible .
3. In step 4 supervised classification approach was adopted.
    - This method should be tested with unsupervised classification approach as well.
    - Simulated Annealing is a lengthy process. The DA based PCM-MRF can be tested with 'gradient descent' method. Also, MPM method has been said to be computationally fast [17], thus, it can also be used.
4. Accuracy Assessment is an important step but in case of sub-pixel classification it is itself a open area of research.
    - The assessment of PCM based method could not be performed directly by using SCM and other composite operator. It is recommended to rescale the output of PCM based methods, as done in case of entropy and then, apply these methods.

5. Other than above points of recommendation, there are more issues that should be looked into.
    - The use of MRF for including spatial contextual information was done in this research work. As one goes towards finer spatial resolution images, the effect of other factors such as texture becomes prominent. Also, contextual information is not limited to spatial domain only. Methods should be exploited to include other contextual information with sub-pixel classification based methods, using MRF
    - If the method to estimate area using the results of the sub-pixel classifier is derived, then this technique can be used to estimate the area occupied by class of interest. E.g. agriculture fields, water body etc.
    - This also gives an eye-opener to study other classification techniques like Linear Mixture Modelling (LMM), SVM etc. with contextual information.

References

[1] Richards J.A & Jia X. ,*Remote Sensing Digital Image Analysis - An Introduction*, 4th edn., Springer-Verlag, Berlin Heidelberg, 2006.

[2] Zhang,J & Foody G M 2001 , "Full-fuzzy classification of sub-urban land cover from remotely sensed imagery : statistical and artificial neural network approaches ",*International Journal of Remote Sensing*, vol 22, no. 4, pp 614-628.

[3] Fisher, P 1997 , "The pixel: a snare and a delusion " ,*International Journal of Remote Sensing*, vol 18, no.3, pp 679-685.

[4] Zadeh, L A 1978 , "Fuzzy set as a basis for theory of possibility",*Fuzzy Sets and Systems*, vol 1, no. 1, pp 3-28.

[5] Foody,G M 2000 , "Estimation of sub-pixel land cover composition in the presence of untrained classes", *Computers & Geosciences*, vol 26,pp 469-478.

[6] Krishnapuram,R & Keller,J M 1993 ,"A possibilistic approach to clustering", *IEEE Transactions on Fuzzy System*, vol 1,no. 2,pp 98-110.

[7] Bezdek,J C ,Ehrlich,R & Full,W 1984 ,"FCM: the fuzzy c-means clustering algorithm', *Computers & Geoscience*, vol 10,no. 2-3,pp 191-203.

[8] Tso,B & Mather,P M 2001 , *Classification Methods for Remotely Sensed Data* , Taylor & Francis, London & New York, 2nd Edition.

[9] Melgani,F & Serpico, S B 2003 , "A Markov Random Field approach to spatio-temporal contextual image classification ",*IEEE Transactions on Geoscience and Remote Sensing*, vol 41, no. 11, pp 2478-2487.

[10] Geman S, and Geman D 1984, "Stochastic Relaxation Gibbs Distribution and the Baeysian Restoration of the Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 6, no.6, pp 721-741.

[11] Krishnapuram, R & Keller, J M 1996, "The Possibilistic C-Means Algorithm: Insights and Recommmendations ", *IEEE Transactions on Fuzzy System*, vol 4, no. 3,pp 385-393.

[12] Pham, D L 2001 , "Spatial Models for Fuzzy Clustering ", *Computer Vision and Image Understanding*, vol 84, pp. 285-297

[13] Dutta, A 2009 , "Fuzzy C-Means classification of multispectral data incorporating spatial contextual information by using Markov Random Field", M.Sc Thesis, GFM, IIRS-ITC JEP.

[14] Mather,P M 1999 , *Computer Processing of Remotely Sensed Images* , J Wiley & Sons,England, 2nd Edition .

[15] Yang,M & Wu,K 2006 , "Unsupervised Possibilistic Clustering ", *Pattern Recognition* ,vol 39, pp 5-21.

[16] Pizurica, A 2002 , "Image Denoising using Wavelets and Spatial Context Modelling", PhD Thesis, Ghentt University.

[17] Solberg,A, Taxt,T & Jain,A K 1996 , "A Markov Random Field Model for Classification of Multisource Satellite Imagery", *IEEE Transactions on Geoscience and Remote Sensing* ,vol 34, no. 1, pp 100-112.

[18] Binaghi,E, Madella,P, Montesano, M.G, & Rampini,A 1997, "Fuzzy Contextual Classification of Multisource Remote Sensing Images", *IEEE Transactions on Geoscience and Remote Sensing* ,vol 35, no. 2, pp 326-340.

[19] Jakkol, 2008(Fall), *Machine Learning, Lecture 19* , MIT Courseware, Massachusetts Institute of Technology, USA (accessed on Feb 08,2009).

[20] Perez,P 1998 , "Markov Random Fields and Images ", *CWI Quarterly* ,vol 11, no. 4, pp 413-437.

[21] Li,S Z 2001 , *Markov Random Field Modelling in Image Analysis*, Springer - Verlag, 2nd Edition.

[22] Li,S Z 1995, "Discontinous MRF prior and robust statistics: a compartive study", *Image and Vision Computing*, vol 13, no. 3,pp 227-233.

[23] Li,S.Z 1995b, "On Discontinuity-Adaptive Smoothness Prior in Computer Vision", *IEEE Transactions on Pattern and Machine Intelligence*, vol 17, no. 6,pp 576-585.

[24] Jackson, Q & Langdrebe,D 2002, "Adaptive Baeysian Contextual Classification Based on Markov Random Fields. ", *Transactions of International Geoscience and Remote Sensing Symposium*, Toronto, 2002.

[25] Okeke,F & Karnieli,A 2006, "Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses. Part I: Algorithm development",*International Journal of Remote Sensing*, vol 27, no. 1-2,pp 153-176.

[26] Schowengerdt,R A 2007, *Remote Sensing : Models and Methods for Image Processing*, Chapter 7, pg 309-311, Academic Press.

[27] "www.nrsa-gov.in/satellite/irs-p6.html", "www.nrsa.gov.in" (accessed on May 18,2009)

[28] Kumar, A, Ghosh, S K & Dadhwal, V K 2007, "Subpixel classifiers: fuzzy theory versus statistical learning algorithm",*SPIE*, vol 27, no. 1-2,pp 153-176.

[29] Kumar, A, Ghosh, S. K & Dadhwal, V. K 2006, "Aster Data processing using statistical learning algorithm ",*Asia-Pacific Remote Sensing*, SPIE symposium, Goa.

[30] Kumar, A, Ghosh, S K & Dadhwal, V K 2007, "Full fuzzy land cover mapping using remote sensing data based on fuzzy c-means and density estimation",*Candadian Journal of Remote Sensing*, vol 33, no. 2,pp 81-87.

[31] Kumar, A, Ghosh, S K & Dadhwal, V K 2007, "A comparision of the performance of fuzzy algorithm versus statistical algorithm based sub-pixel classifier for remote sensing data",*ISPRS Commission VII Mid-term Symposium "Remote Sensing : From Pixels to Processes"* Enscheded, Netherlands, 8-11 May 2006.

[32] Tso,B & Olsen,R.C 2005, "A contextual classification scheme based on MRF model with improved parameter estimation and multi-scale fuzzy line process,", *Remote Sensing of Environment*, vol 97, no. 2, pp 127-136.

[33] Reulke,R & Lippock,A 2008, "Markov Random Field based texture segmentation for road extraction,", *The International Archives of Photogrammetry, Remote Sensing and Spatial information Sciences*, vol XXXVII, part B3b, Beijing.

[34] Smits,P C & Dellepiane,S G 1997, "Discontinuity Adaptive MRF Model for the Analysis of Synthetic Aperture Radar Images,", *IEEE Transactions of Geoscience and Remote Sensing*, vol 35, no. 4, pp 837-840.

[35] Hou,Z ,Qian,W, Huang,S, Hu,Q, & Nowinski, W.L, 2007, "Regularized fuzzy c-means method for brain tissue clustering,", *Pattern Recognition Letters*, vol 28, no. 13, pp 1788-1794.

[36] Smits,P C & Dellepiane,S G 1997, "Synthetic Aperture Radar Image Segmentation by a detailed preserving Markov Random Field Approach,", *IEEE Transactions of Geoscience and Remote Sensing*, vol 35, no. 4, pp 844-857.

[37] Kirkpatrick,S, Gelett,C D & Vecchi,M P 1983, "Optimization by simulated annealing,", *Science*, vol 220, pp 621-630.

[38] Cerny,V 1985, "A thermodynamic approach to the traveling salesman problem: An efficient simulation", *Journal of Optimization Theory and Applications*, vol 45, pp 41-51.

[39] Bertsimas,D & Tsitsiklis, J 1993, "Simulated Annealing", *Statistical Science*, vol 8, no. 1,pp 10-15.

[40] Walsh,B 2004, *Markov Chain Monte Carlo and Gibbs Sampling* , Lecture Notes for EEB 581,MIT Courseware, Massachusetts Institute of Technology,USA (accessed on Sep 1,2009).

[41] Ricotta,C 2004, "Evaluationg the classificaiton accuracy of fuzzy thematic maps with a simple parameteric measure", *International Journal of Remote Sensing*, vol 25, no. 11,pp 2169-2176.

[42] Binaghi, E, Brivio, P A , Ghezzi,P, & Rampini, A. 1999, "A fuzzy set-based accuracy assessment of soft classification", *Pattern Recognition Letters*, vol 20, pp 935-948.

[43] Silvan-Cardenas, J L & Wang,L 2008, "Sub-pixel confusion uncertainty matrix for assessing the soft classifications", *Remote Sensing of Environment*, vol 112, pp 1081-1095.

[44] Smith,S.W 1997 , *DSP Guide for Scientists and Engineers*, chapter 2,California Technical Publications.

[45] DeCoursey,W J 2003 , *Statistics and Probability for Engineering Applications*, Elseiver Science, USA.

[46] Pontius Jr,R G & Cheuk,M L 2006, "A generalized cross tabulation marix to compare soft classified maps at multiple resolutions", *International Journal of Geographical Information Science*, vol 20, no. 1,pp 1-30.

[47] Dehghan,H & Ghassemian,H 2006, "Measurement of uncertainty by the entropy: application to the classification of MSS data", *International Journal of Remote Sensing*, vol 27, no. 18,pp 4005-4014.

[48] Ricotta,C & Avena, G C 2006, "Evaluating the degree of fuzziness of thematic maps with a generalized entropy function : a methodological outlook", *International Journal of Remote Sensing*, vol 23, no. 20, pp 4519-4523.

[49] Kitaw,H G 2007 ,"Image Pan-sharpening with Markov Random Field and Simulated Annealing", M.Sc Thesis, GFM, International Institute of Geoinformation Science and Earth Observation.

[50] Kassaye,R H 2006 ,"Suitability of Markov Random Field-based Method for Super-Resolution Mapping ", M.Sc Thesis, GFM, International Institute of Geoinformation Science and Earth Observation.

[51] Kumar, A, Ghosh, S K, & Dadhwal, V K 2006, "Sub-pixel land cover mapping: SMIC system.", *IIn proceedings of ISPRS TC-IV*, Internationl Symposium on Geospatial Database for Sustainable Development, held at Goa from Sep 27-30, ISPRS Archives, Enschede, The Netherlands, 36(4): pp.97-101.

[52] Wilkinson,D.J 2004, *Lecture Notes for MAS451,http://www.staff.ncl.ac.uk/d.j.wilkinson.html*, New Castle University,UK (accessed on Sep 7,2009).

[53] Jensen, J.R 1996 , *Introductory Digital Image Processing*, Prentice Hall, New Jersey, 2nd Edition,.

[54] Tolpekin, V A & Stein,A 2009, "Quantification of the effects of the land-cover-class seperability on the accuracy of Markov-Random-Field-Based superresolution mapping", *IEEE Transactions on Geoscience and Remote Sensing*, vol 23,no. 20,pp 4519-4523.

Appendix A

A.1 Accuracy Assessment results of AWIFS and LISS III images classified using FCM

Table A.1: Results of AWIFS images clasified using FCM vs. Reference data set generated using LISS IV

Accuracy Assessment methods –¿	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	94.52	93.38
Producer's Accuracy	74.315	63.56
Agriculutral fields with crop		
User's Accuracy	77.52	75.21
Producer's Accuracy	59.58	54.01
Dry agricultural field without crop		
User's Accuracy	53.99	34.13
Producer's Accuracy	89.62	87.74
Moist agricultural field without crop		
User's Accuracy	72.96	61.69
Producer's Accuracy	76.97	73.24
Eucalyptus Plantation		
User's Accuracy	69.42	66.55
Producer's Accuracy	89.23	87.55
Broad Leaved Forest		
User's Accuracy	90.57	87.98
Producer's Accuracy	69.33	62.69
Average User's Accuracy	76.50	
Average Producer's Accuracy	76.51	
Fuzzy Overall Accuracy	76.02	70.67
Fuzzy Kappa Coefficient		0.621

Table A.2: Results of LISS III images clasified using FCM vs. Reference data set generated using LISS IV

Accuracy Assessment methods	**FERM (%)**	**MIN-MIN (%)**
Water		
User's Accuracy	96.16	95.82
Producer's Accuracy	88.53	82.58
Agriculutral fields with crop		
User's Accuracy	84.44	80.68
Producer's Accuracy	88.89	86.73
Dry agricultural field without crop		
User's Accuracy	87.60	80.39
Producer's Accuracy	82.51	73.63
Moist agricultural field without crop		
User's Accuracy	77.71	73.24
Producer's Accuracy	87.24	82.84
Eucalyptus Plantation		
User's Accuracy	91.78	90.63
Producer's Accuracy	71.77	75.213
Broad Leaved Forest		
User's Accuracy	76.12	71.69
Producer's Accuracy	93.25	92.89
Average User's Accuracy	84.63	
Average Producer's Accuracy	85.36	
Fuzzy Overall Accuracy	84.64	82.59
Fuzzy Kappa Coefficient		0.7804

A.2 Results of fuzzy kappa coefficients obtained using SCM, MIN-PROD and MIN-LEAST on PCM classified data

Table A.3: The table shows absurd results obtained using SCM, MIN-PROD & MIN-LEAST

Accuracy Assessment methods	**SCM (%)**	**MIN-PROD (%)**	**MIN-LEAST**
Fuzzy Kappa Coefficient	0.330	1.034	-0.131

A.3 Output fraction images after classifying AWIFS and LISS III with PCM-MRF II

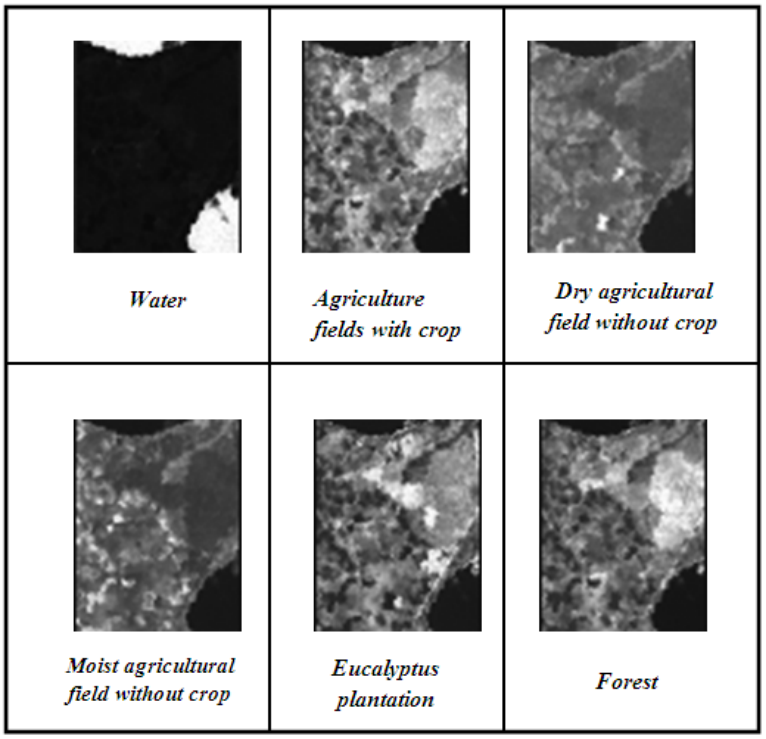


Figure A.1: Resultant fraction images on classifying AWIFS with PCM-MRF II

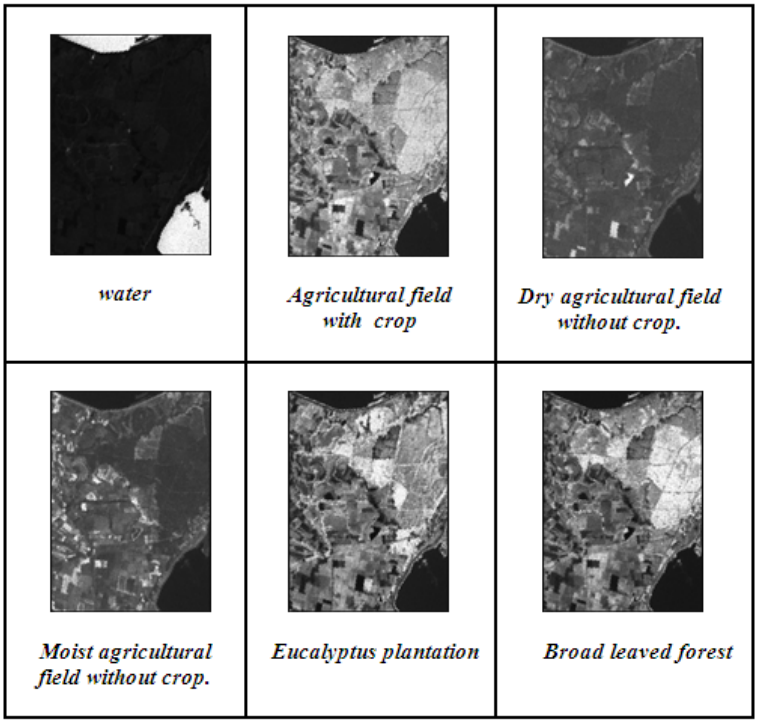


Figure A.2: Resultant fraction images on classifying LISS III with PCM-MRF II

A.4 Snapshots of SMIC : In-house Tool developed by Kumar et al

source : Kumar et al, ISRS 2008 symposium

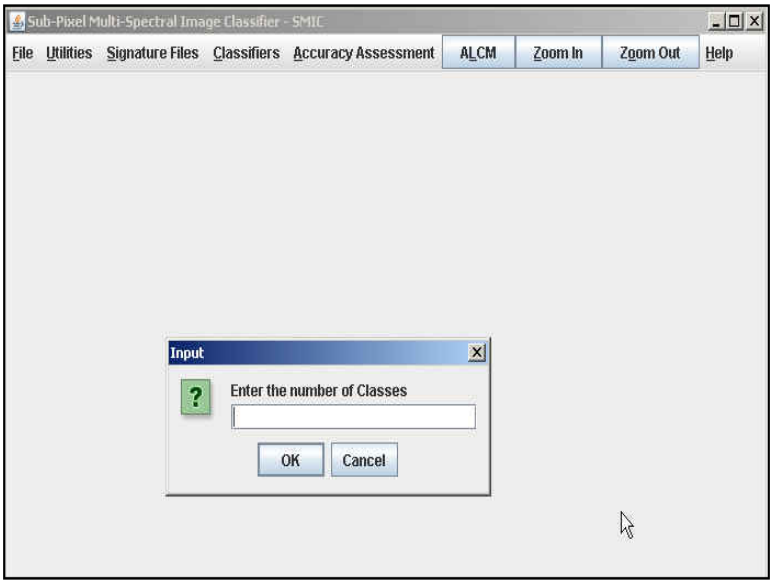


Figure A.3: Snapshot I of the in-house tool

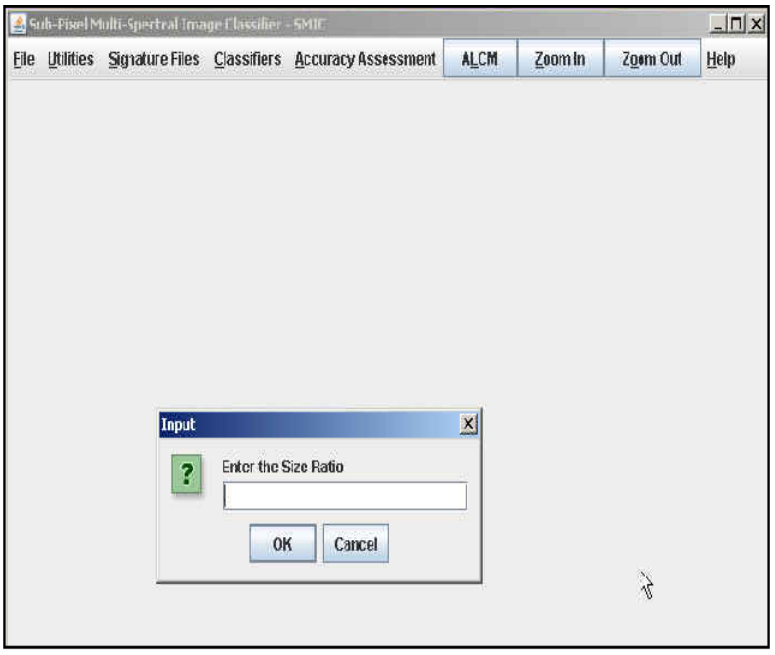


Figure A.4: Spatial Resolution ratio between classified and reference sub-pixel

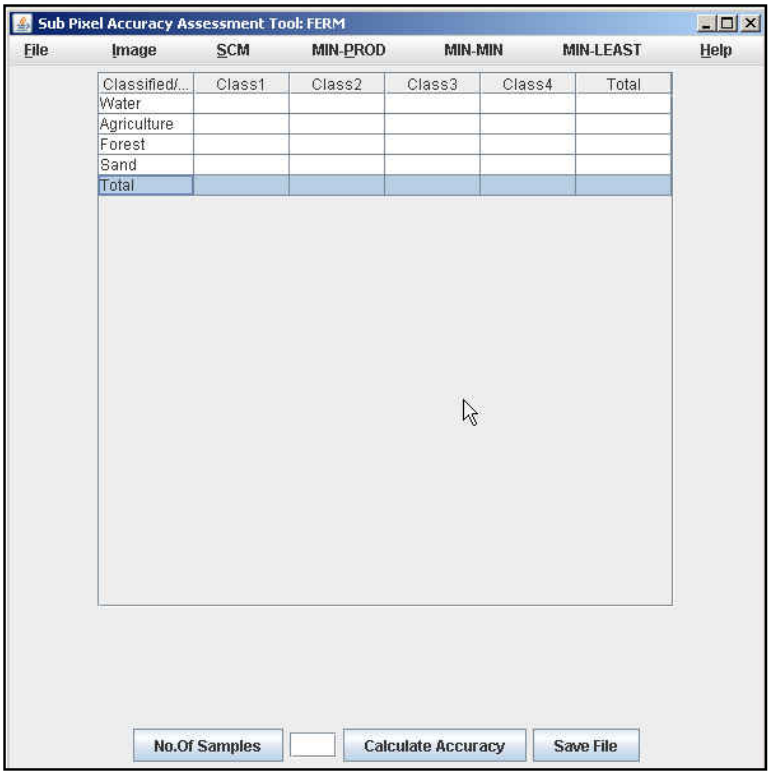


Figure A.5: Fuzzy error matrices using various assessment methods

Appendix B

Publication

The methods developed during this research have been accepted/submitted in the following symposiums :-

1. Paper accepted for Oral Presentation at ISRS Symposium -2009, held at Nagpur (Sept 17-19, 2009) "Parameter Estimation for Possibilistic c-Means spatial contextual information based sub-pixel classification for multispectral images", Sivee Chawla, Anil Kumar, Valentyn Tolpekin, Nicholas H Hamm.
2. Paper submitted at IGARSS 2010 "Markov Random Field and Standard Regularization model based Possibilistic - c -Means soft classification approach." Sivee Chawla, Anil Kumar, Valentyn Tolpekin, Nicholas H Hamm.