

MODULE – 5 LECTURE NOTES – 3

FUZZY CLASSIFICATION

1. Introduction

Hard classification is based on classical set theory in which precisely defined boundaries are generated for a pixel as either belonging to a particular class or not belonging to that class. During hard classification, each individual pixel within a remotely sensed imagery is given a class label. This technique works efficiently when the area imaged is homogeneous in nature. But geographical information is heterogeneous in nature. This implies that the boundaries between different land cover classes are fuzzy that gradually blend into one another. Fuzziness and hardness are characteristics of landscape at a particular scale of observation. If the aim of end user is to label each pixel unambiguously, the existence of heterogeneous pixels containing more than one land cover type will create a problem. This is owing to the fact that the pixel may not fall clearly into one of the available classes as it represents mixed classes. This problem also surfaces if the satellite borne instrument imaging earth has a large field of view (1 km or more). Fuzzy set theory provides useful concepts to work with imprecise data. Fuzzy logic can be used to discriminate among land cover types using membership functions. These are elaborated in the sections below.

2. Fuzzy Set Theory

Researchers in the field of psycholinguistics have investigated the way humans evaluate concepts and derive decisions. Analysis of this kind of uncertainty usually results in a perceived probability rather than the mathematically defined mobility, which forms the basis of fuzzy sets (Zadeh, 1973). The theory of fuzzy sets was first introduced when it was realized that it may not be possible to model ill-defined systems with precise mathematical assumptions of the classical methods, such as probability theory (Chi *et al.*, 1996). The underlying logic of fuzzy-set theory is that it allows an event to belong to more than one sample space where sharp boundaries between spaces are hardly found.

If X be a collection of objects with elements noted as x , then a fuzzy set A in X is a set of ordered pairs given by the underlying relation (Equation (2.1)):

$$A = \left\{ \frac{\mu_A(x)}{x} \right\} \quad (1)$$

Where $\mu_A(x)$ is the characteristic function or membership grade of x in A which may take any real value in the interval $[0, 1]$. Thus, the nearer the value of $\mu_A(x)$ to unity, the higher the grade of membership of x in A . Fuzzy set theory is not a theory that permits vagueness in our computations, but it is rather a methodology to show how to tackle uncertainty, and to handle imprecise information in a complex situation. The fuzzy models which are applicable in pattern recognition and image processing, involve the use of fuzzy membership function, fuzzy clustering, fuzzy-rule based systems, fuzzy entropy, fuzzy integrals, etc.

The operations on fuzzy sets presented in this section are based on the original works of Zadeh (1965) and these should not be considered as a complete collection.

1. *Fuzzy Union*: The union of two fuzzy sets A and B with respective membership functions $\mu_A(x)$ and $\mu_B(x)$ is a fuzzy set C , written as $C = A \cup B$, whose membership value is the smallest fuzzy set containing both A and B .

$$\mu_C(x) = \text{Max}[\mu_A(x), \mu_B(x)] , \quad x \in X \quad (2)$$

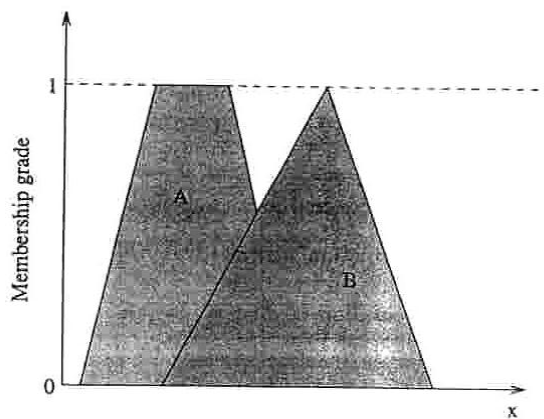


Figure 1 Fuzzy union of two sets A and B

2. *Fuzzy Intersection*: The intersection of two fuzzy sets A and B with respective membership functions $\mu_A(x)$ and $\mu_B(x)$ is a fuzzy set C , written as $C = A \cap B$ whose membership function is related to those of A and B by:

$$\mu_C(x) = \text{Min}[\mu_A(x), \mu_B(x)], \quad x \in X \quad (3)$$

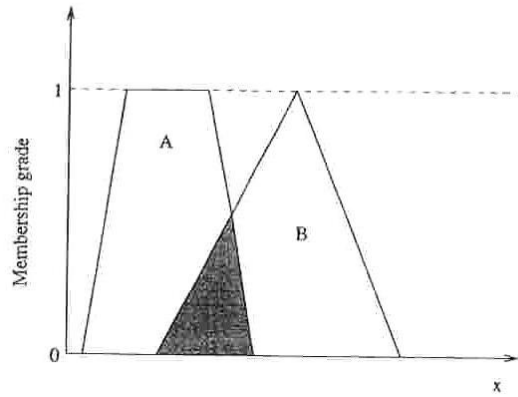


Figure2 Fuzzy intersection of two sets A and B

3. *Fuzzy complement*: The complement of a fuzzy set A is denoted by A' and is defined by

$$\mu_{A'} = 1 - \mu_A \quad (4)$$

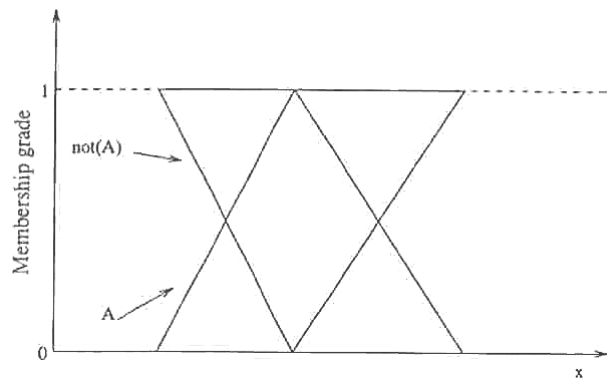


Figure 3: Fuzzy complement

With the operations of union, intersection and complementation, it is easy to extend many of the basic identities which hold good for ordinary sets to fuzzy sets (i.e. De Morgan's laws, distributive laws, etc). In addition to these operators, one can define a number of other ways of forming combinations of fuzzy sets and relating them to one another. These are summation, product, difference, convex combination of fuzzy sets.

3. Membership Function

The membership function is the underlying power of every fuzzy model as it is capable of modeling the gradual transition from a less distinct region to another in a subtle way (Chi *et al.*, 1996). Membership functions characterize the fuzziness in a fuzzy set, whether the elements in the set are discrete or continuous, in a graphical form for eventual use in the mathematical formalisms of fuzzy set theory. But the shapes used to describe fuzziness have very few restrictions indeed. There are an infinite number of ways to graphically depict the membership functions that describe this fuzziness (Ross *et al.*, 2002). Since membership functions essentially embody all fuzziness for a particular fuzzy set, its description is the essence of a fuzzy property or operation. Because of the importance of the shape of the membership function, a great deal of attention has been focused on development of these functions. There are several standard shapes available for membership functions like triangular, trapezoidal and Gaussian, etc. The direct use of available shapes for membership function is found effective for image enhancement where different types of membership functions are used to reduce the amount of iterations carried out by a relaxation technique and provides a better way to handle the uncertainty of the image histogram. The choice of membership function is problem dependent which requires expert knowledge (Zadeh, 1996). In situations wherein prior information about data variation is not available, membership values can also be generated from the available data using clustering algorithms, which is the normal practice.

4. Fuzzy classification methods

Fuzzy set theory provides useful concepts and tools to deal with imprecise information and partial membership allows that the information about more complex situations such as cover mixture or intermediate conditions be better represented and utilized (Wang, 1990). Use of fuzzy sets for partitioning of spectral space involves determining the membership grades attached to each pixel with respect to every class. Instead of being assigned to a single class, out of m possible classes, each pixel in fuzzy classification has m membership grade values, where each pixel is associated with a probability of belonging to each of the m classes of interest (Kumar, 2007). The membership grades may be chosen heuristically or subjectively. Heuristically chosen membership functions do not reflect the actual data distribution in the input and the output spaces. Another option is to build membership functions from the data

available for which, we can use a clustering technique to partition the data, and then generate membership functions from the resulting clusters. A number of classification methods may be used to classify remote sensing image into various land cover types. These methods may be broadly grouped as supervised and unsupervised (Swain and Davis, 1978). In fuzzy unsupervised classification, membership functions are obtained from clustering algorithms like C-means or ISODATA method. In fuzzy supervised classification, these are generated from training data.

The classification of remotely sensed imagery relies on the assumptions that the study area is composed of a number of unique, internally homogeneous classes, classification analysis is based on reflectance data and that ancillary data can be used to identify these unique classes with the aid of ground data (Lillesand and Kiefer, 1994). The fuzzy approaches are adopted as they take into account the fuzziness that may be characteristic of the ground data (Foody, 1995). Zhang and Foody (1998) investigated fuzzy approach for land cover classification and suggested that fully fuzzy approach holds advantages over both the conventional hard methods and partially fuzzy approaches. Clustering algorithms can be loosely categorized by the principle (objective function, graph-theoretical, hierarchical) or by the model type (deterministic, statistical and fuzzy). In the literature on soft classification, the fuzzy c-mean (FCM) algorithm is the most popular method (Bastin 1997; Wu and Yand 2002; Yang *et al.*, 2003). One of the popular parametric classifiers based on statistical theory is the Fuzzy Gaussian Maximum Likelihood (FGML) classifier. This is an extension of traditional crisp maximum likelihood classification wherein, the partition of spectral space is based on the principles of classical set theory. In this method, land cover classes can be represented as fuzzy sets by the generation of fuzzy parameters from the training data. Fuzzy representation of geographical information makes it possible to calculate statistical parameters which are closer to the real ones. This can be achieved by means of the probability measures of fuzzy events (Zadeh, 1968). Compared with the conventional methods, this method has proved to improve remote sensing image classification in the aspects of geographical information representation, partitioning of spectral space, and estimation of classification parameters (Wang, 1990). Despite the limitations due to its assumption of normal distribution of class signature (Swain and Davis, 1978), it is perhaps one of the most widely used classifiers (Wang, 1990; Hansen *et al.*, 1996; Kumar, 2007).

A hard classification fails to recognize or to represent the existence of classes and objects which grade into one another. Such simplification can be seen as a waste of the available multi-spectral information, which could be more efficiently interpreted (Wang, 1990; Foody, 1995). An alternative to this problem is the use of advanced methods of image classification like sub-pixel classification approaches. Some of the fuzzy classification methods adopted in this research work are explained in the following sub sections.

In remote sensing, pixel measurement vectors are often, considered as points in a spectral space. Pixels with similar spectral characteristics form groups which correspond to various ground-cover classes that the analyst defines. The groups of pixels are referred to as spectral classes, while the cover classes are information classes. To classify pixels into groups, the spectral space should be partitioned into regions, each of which corresponds to one of the information classes defined. Decision surfaces are defined precisely by some decision rules (for example, the decision rule of conventional maximum likelihood classifier) to separate the regions. Pixels inside a region are classified into the corresponding information class. Such a partition is usually called a hard partition. Fig 4.1a illustrates a hard partition of spectral space and decision surfaces. A serious drawback of the hard partition is that a great quantity of spectral information is lost in determining the pixel membership, Let X be a universe of discourse; whose generic elements are denoted x : $X = \{x\}$. Membership in a classical set A of X is often viewed as a characteristic function χ_A from $\{0,1\}$ such that $\chi_A(x) = 1$ if and only if $x \in A$. A *fuzzy set* (Zadeh, 1965) B in X is characterized by a *membership function*, f_B , which associates with each x a real number in $[0,1]$. $f_B(x)$ represents the "grade of membership" of x in B . The closer the value of $f_B(x)$ is to 1, the more x belongs to B . A fuzzy set does not have sharply defined boundaries and an element may have partial and multiple memberships.

Fuzzy representation of geographical information enables a new method for spectral space partition. When information classes can be represented as fuzzy sets, so can the corresponding spectral classes. Thus a spectral space is not partitioned by sharp surfaces. A pixel may belong to a class to some extent and at the same time belong to another class to some other extent. Membership grades are attached to indicate these extents. Such a partition is referred to as a fuzzy partition of spectral space. Fig 1b illustrates membership grades of a pixel in a fuzzy partition. A fuzzy partition of spectral space can represent a real situation better than a hard partition and allows more spectral information to be utilized in subsequent

analysis. Membership grades can be used to describe cover class mixture and intermediate cases.

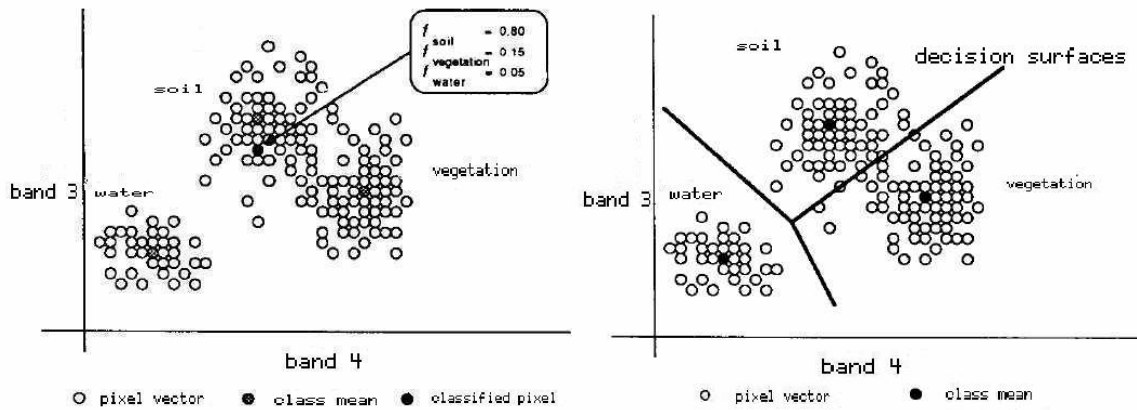


Figure 4 Hard partition of spectral space and decision surfaces; 1b: membership grades of a pixel in fuzzy partition of spectral space.

4.1 Supervised Approach

This algorithm is an extension of the conventional hard per pixel Gaussian Maximum Likelihood (GML) algorithm. Fuzzy representation of geographical information enables a new method for spectral space partition. In this approach, information classes as well as the spectral classes are represented as fuzzy sets. Thus, a spectral space is not partitioned by sharp boundaries. A fuzzy partition of spectral space/fuzzy partition matrix (μ) is a family of fuzzy sets F_1, F_2, \dots, F_m on universe X , Where F_1, F_2, \dots, F_m represent the spectral classes, c is the number of predefined classes, x is a pixel measurement vector and μ_{F_1} is the membership function.

The fuzzy partition matrix can be recorded as:

$$\mu_{F(x)} = \begin{bmatrix} \mu_{F_1(x_1)} & \mu_{F_1(x_2)} & \cdots & \mu_{F_1(x_N)} \\ \mu_{F_2(x_1)} & \mu_{F_2(x_2)} & & \\ \cdot & \cdot & \cdots & \\ \mu_{F_c(x_1)} & \mu_{F_c(x_2)} & \cdots & \mu_{F_c(x_N)} \end{bmatrix}$$

Membership grades can be used to describe cover class mixture and intermediate classes. In the process, the stray pixels between classes may be classified as such. In Supervised Approach, which is similar to maximum likelihood classification approach, instead of normal mean vector and covariance matrices, fuzzy mean vectors and fuzzy covariance matrices are developed from statistically weighted training data, and the training areas may be a combination of pure and mixed pixels. By knowing mixtures of various features, the fuzzy training class weights are defined. A classified pixel is assigned a membership grade with respect to its membership in each information class. In this procedure the conventional mean and covariance parameters of training data are represented as a fuzzy set. The following two equations (1, 2) describe the fuzzy parameters of the training data:

$$M_c^* = \frac{\sum_{i=1}^n \mu_c(x_i) x_i}{\sum_{i=1}^n \mu_c(x_i)} \quad (5)$$

$$V_c^* = \frac{\sum_{i=1}^n \mu_c(x_i) (x_i - M_c^*) (x_i - M_c^*)^T}{\sum_{i=1}^n \mu_c(x_i)} \quad (6)$$

where, M_c^* is the fuzzy mean of training class c ; V_c^* is the fuzzy covariance of training class c ; x_i is the vector value of pixel i , $\mu_c(x_i)$ is the membership of pixel x_i to training class c , n is the total number of pixels in the training data. In order to find the fuzzy mean (eqn. 1) and fuzzy covariance (eqn. 2) of every training class, the membership of pixel x_i to the training class c must be first known. Membership function to class c based on the conventional maximum likelihood classification algorithm with fuzzy mean and fuzzy covariance is:

$$\mu_c(x_i) = \frac{P_c^*(x_i)}{\sum_{j=1}^m P_j^*(x_i)} \quad (7)$$

where, $P_c^*(x_i)$ is the maximum likelihood probability of pixel x_i to class c , m is the number of classes. The membership grades of a pixel vector x depend upon the pixel's position in the

spectral space. The *a posteriori* probabilities are used to determine the class proportions in a pixel. The algorithm iterates until there is no significant change in the membership values obtained.

As this method is an extension of MLC, it inherits its advantages and disadvantages. The disadvantage is the normality of data assumption which it is based upon. Compared with the conventional methods, FGML improves remote sensing image classification in the aspects of: 1) Representation of geographical information, 2) Partitioning of spectral space, and 3) Estimation of classification parameters (Wang, 1990).

4.2 Unsupervised Approach

4.2.1 Fuzzy C-means (FCM)

Fuzzy C-means clustering also known as Fuzzy ISODATA is an iterative technique which is separated from hard c-means that employ hard partitioning. The FCM employs fuzzy partitioning such that a data point can belong to all groups with different membership grades between 0 and 1. The aim of FCM is to find cluster centroids that minimize the dissimilarity function. Differing from hard clustering techniques such as c-means, which will converge the objective function iteratively to a local minimum from each sample to the nearest cluster centroid, fuzzy clustering methods assign each training sample a degree of uncertainty described by a membership grade. A pixel's membership grade function with respect to a specific cluster indicates to what extent its properties belong to that cluster. The larger the membership grade (close to 1), the more likely that the pixel belongs to that cluster. FCM algorithm was first introduced by Dunn (1973); and the related formulation and the algorithm was extended by Bezdek (1974). The purpose of FCM approach, like the conventional clustering techniques; is to minimize the criteria in the least squared error sense. For $c \geq 2$ and m any real number greater than 1, the algorithm chooses $\mu_i, : X \rightarrow [0,1]$ so that $\sum_i \mu_i = 1$ and $w_j \in R^d$ for $i=1,2,...,c$ to minimize objective function

$$J_{FCM} = \frac{1}{2} \sum_{j=1}^c \sum_{i=1}^n (\mu_{i,j})^m \|x_i - w_j\|^2 \quad (8)$$

where $\mu_{i,j}$ is the value of the j^{th} membership grade on the i^{th} sample x_i . The vectors $w_1, ..., w_j, ..., w_c$, called cluster centroids, can be regarded as prototypes for clusters represented by the membership grades. For the purpose of minimizing the objective function, the cluster

centroids and membership grades are chosen so that a high degree of membership occurs for samples close to the corresponding centroids. The FCM algorithm, a well-known and powerful method in clustering analysis, is further modified as follows.

4.2.2 Penalized Fuzzy C-Means (PFCM) Algorithm

Another strategy of the fuzzy clustering method, called penalized fuzzy c-means (PFCM) algorithm with the addition of a penalty term, was demonstrated by Yang (1993) and Yang and Su (1994). It is an FCM algorithm of generalized type depending upon the penalized term in accordance with the value of w . It was shown by Yang (1993) that the PFCM algorithm is more meaningful and effective than the FCM method.

$$J_{PFCM} = \frac{1}{2} \sum_{j=1}^c \sum_{i=1}^n (\mu_{i,j})^m \|x_i - x_j\|^2 - \frac{1}{2} \nu \sum_{j=1}^c \sum_{i=1}^n \mu_{i,j}^m \ln \alpha_j \quad (9)$$

where α_j is a proportional constant the value of class j and $\nu (\geq 0)$ is a constant. When $\nu = 0$, J_{PFCM} equals J_{FCM} . The penalty term is added to the J_{FCM} objective function, where

$$\alpha_j = \frac{\sum_{i=1}^n \mu_{i,j}^m}{\sum_{j=1}^c \sum_{i=1}^n \mu_{i,j}^m}, \quad j = 1, 2, \dots, c \quad w_j = \frac{\sum_{i=1}^n \mu_{i,j}^m x_i}{\sum_{i=1}^n \mu_{i,j}^m} \quad (10)$$

$$\mu_{i,j} = \left(\frac{\sum_{l=1}^c \left(\|x_i - w_l\|^2 - \nu \ln \alpha_l \right)^{1/(m-1)}}{\left(\|x_i - w_j\|^2 - \nu \ln \alpha_j \right)^{1/(m-1)}} \right)^{-1}; \quad i = 1, 2, \dots, n; j = 1, 2, \dots, c \quad (11)$$

PFCM algorithm is presented as follows:

1. Randomly set cluster centroids w_j ($2 \leq j \leq c$), fuzzification parameter m ($1 \leq m \leq \infty$), and the value $\varepsilon > 0$. Give a fuzzy c-partition $U^{(0)}$.
2. Compute $\alpha_j^{(t)}$, $w_j^{(t)}$, with $U^{(t-1)}$ using equation (6). Calculate the membership matrix $U = [\mu_{i,j}]$ with $\alpha_j^{(t)}$, $w_j^{(t)}$ using equation (7).
3. Compute $\Delta = \max(|U^{(t-1)} - U^{(t)}|)$. If $\Delta > \varepsilon$, then go to step 2; otherwise go to step 4.

4. Find the results for the final class centroids.

In the last step, a defuzzification process should be applied to the fuzzy partition data to obtain the final segmentation. A pixel is assigned to a cluster when its membership grade in that cluster is the highest. The disadvantage of FCM is that due to the use of an inner-product norm induced distance matrix, its performance is good only when the data set contains clusters of roughly the same size and shape. Also, since it is unsupervised, the order of occurrence of class fraction images cannot be predicted. However, the independence of this algorithm to any type of data distribution makes it popular among all the clustering algorithms.

5. Assessing classification accuracy

Evaluation of classification results is an important process in the classification procedure. Traditionally, the accuracy is determined empirically by comparing with corresponding reference or ground data wherein the results are tabulated in the form of a square matrix known as confusion matrix (Card, 1982). Ideal situation is represented by a diagonal matrix where only principal diagonal elements have non-zero values *i.e.* all areas of the image have been correctly classified (Genderen and Lock, 1977). Ideally, classification accuracy should be expressed in the form of a single index which is readily interpretable and which allows the relative performance of different classifications to be evaluated (Foody, 1994). For evaluating the output of soft classification, several methods have been proposed. Recently, a novel method of assessing fuzzy classification accuracy using the concept of fuzzy sets was introduced called Fuzzy error matrix (FERM) (Binaghi, 1999). The FERM generated using fuzzy set theory is an extended version of the traditional error matrix. To assess the accuracy of soft classification, gradual membership in several classes is allowed for each element of sample data and assignments to classes are judged correct, or incorrect in varying degrees. Thus, the accuracy assessment may best be described as measuring the degree of agreement or correspondence to the ground data and are not necessarily a true reflection of the closeness to reality (Foody, 2002). The quality of clustering is indicated by the closeness with which data points are associated to the cluster centers. This can be determined by calculating various cluster validity functions like partition coefficient, classification entropy, etc (Michael, 1982).

5.1 Fuzzy Error Matrix (FERM)

Accuracy of classifier is generally assessed empirically by selecting a sample of reference data and comparing their actual class assignments with those provided by the automated classifier. The common technique is to represent the information in the form of confusion matrix. Binaghi (1999) was the first to extend the applicability of the traditional error matrix to the evaluation of soft classifiers using fuzzy set theory. Fuzzy error matrix is a square array of integer numbers set out in rows and columns that represent the number of samples of the actual category assigned to a particular class. The diagonal elements show the number of sample elements which have been classified correctly, while the off diagonal elements show the number of samples incorrectly classified.

Let R_n be the set of fuzzy reference data assigned to class n , C_m be the set of fuzzy classification data assigned to class m , $\mu_{R_n}(x)$ and $\mu_{C_m}(x)$ represent the gradual membership of the sample elements in classes n and m . The fuzzy set operators can be used within the matrix building procedure to provide a fuzzy error matrix M . The assignment to the element $M(m,n)$ involves the computation of the degree of membership in the fuzzy intersection set $C_m \cap R_n$. The ‘min’ operator is introduced in determination of values of the fuzzy error matrix (Equation (3.24):

$$M(m,n) = |C_m \cap R_n| = \sum_{x \in X} \mu_{C_m \cap R_n}(x) \quad (12)$$

The layout of a typical FERM is shown in Table 5.1

Table 5.1 General layout of Fuzzy error matrix

Fuzzy classification	Reference classification					User's Accuracy
	Class 1	Class 2	Class c	
Class 1	$M_{(1,1)}$	$M_{(1,2)}$	$M_{(1,c)}$	
Class 2	$M_{(2,1)}$	$M_{(2,2)}$	$M_{(2,c)}$	
...	
...	
Class c	$M_{(c,1)}$	$M_{(c,2)}$	$M_{(c,c)}$	
Producer's Accuracy						Overall Accuracy

The fuzzy error matrix can be used as the starting point for descriptive techniques in the same manner as used in the conventional error matrix.

(a) User Accuracy

It is a measure of commission error and is calculated by dividing the corresponding element of the major diagonal by the total of grades of membership found in the column marginal in reference and classification data as given in Equation (3.25).

$$\text{User Accuracy} = \frac{X_{jj}}{X_{+j}} \quad (13)$$

Where, X_{jj} = diagonal element at j^{th} column and j^{th} row.

X_{+j} = column marginal total in j^{th} column of confusion matrix.

(b) Producer Accuracy

It represents a measure of the omission error and is defined as the ratio of diagonal membership totals to the total of membership grades found in the row marginal in classification and reference data as shown in Equation (3.26):

$$\text{Producer Accuracy} = \frac{X_{ii}}{X_{i+}} \quad (14)$$

Where, X_{ii} = diagonal element at i^{th} column and i^{th} row.

X_{i+} = row marginal total in i^{th} row of confusion matrix.

(c) Overall Accuracy

This is the simplest index which gives the total match between reference and classification data. It is given by Equation (3.27) as :

$$\text{Overall Accuracy} = \frac{\sum X_{ii}}{\sum X_{ij}} \quad (15)$$

$\sum X_{ii}$ is total grades of membership found in reference and $\sum X_{ij}$ is the sum of diagonal elements.

(d) Kappa Co-efficient

The proportion of chance agreement is defined as per Equation (3.28) as:

$$p_c = \frac{(\sum X_{+i} X_{i+})}{N^2} \quad (16)$$

Where, the quantities X_{+i} and X_{i+} represent column marginal and row marginal total of membership grades respectively. Using this value of chance agreement Kappa Co-efficient or *Khat* index is defined using Equation (3.29) (Stein, Meer and Gorte, 2002):

$$\kappa = \frac{(p_0 - p_c)}{(1 - p_c)} \quad (17)$$

Where p_0 is the overall accuracy obtained from the confusion matrix.

(e) Z statistic

This test determines if the results of two error matrices are statistically similar or not (Congalton *et al.*, 1983). It is calculated using Equation (3.30)

$$Z = \frac{\kappa_a - \kappa_b}{\sqrt{\sigma_a^2 + \sigma_b^2}} \quad (18)$$

where Z is the test statistic for significant difference in large samples, κ_a and κ_b are the *Khat* indices for two error matrices a and b with variance for *Khat* indices as σ_a^2 and σ_b^2 .

6. Case Study

Fuzzy clustering algorithms explained in the previous section are applied to identify Paddy, Semi-dry and Sugarcane crops using IRS LISS I (Linear Imaging Self Scanner) data in the Bhadra command area for Rabi season of 1993. Bhadra dam is located in Chickmagalur District of Karnataka state. The dam is situated 50 km upstream of the point where Bhadra river joins Tunga, another tributary of Krishna river, and intercepts a catchment of almost 2000 sq.km. Bhadra reservoir system consists of a storage reservoir with a capacity of 2025 M m³, a left bank canal and a right bank canal with irrigable areas of 7,031 ha and 92,360 ha respectively. Figure 6.1 shows location map of Bhadra command area. Major crops cultivated in the command area are Paddy, Semi-dry and Sugarcane. Paddy transplantation is staggered over a period of more than a month and semi-dry crops are sown considerably earlier to Paddy. The command area is divided into three administrative divisions, viz., Bhadravati, Malebennur and Davangere.

Satellite imageries used for the study are acquired from IRS LISS I (with spatial resolutions of 72.5 m) on dates 20th February, 14th March and 16th April in the years 1993. Figure 6.2 shows the standard FCC (False colour composite) of Bhadra command area on 16th April 1993. For the study area, ground truth was collected for various crops by scientists from National Remote Sensing Agency, Hyderabad, during Rabi 1993, by visiting the field.

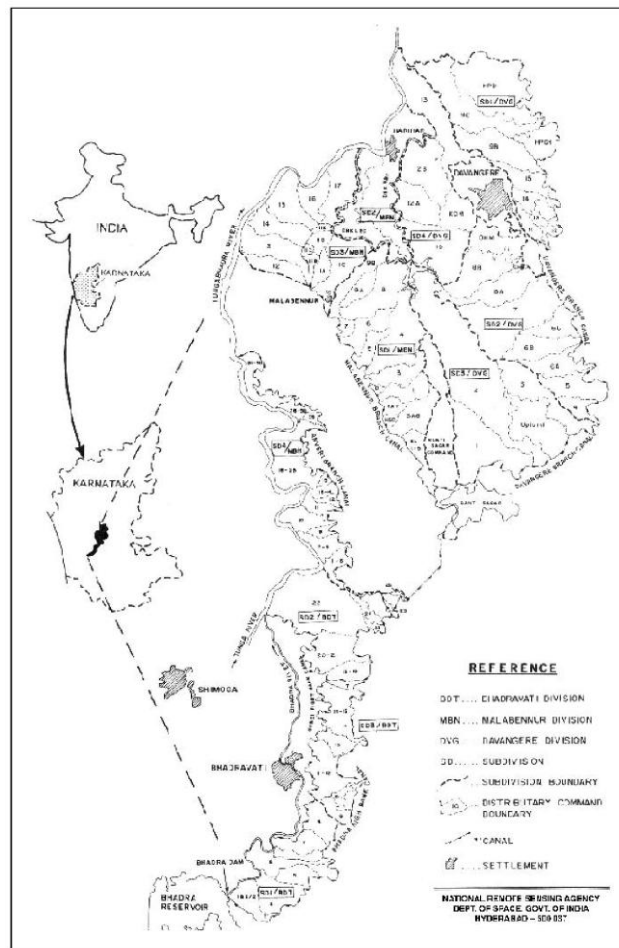


Figure 5 Location map of Bhandra command area

6.1 Results and Discussions

Use of penalized fuzzy c-means algorithm requires selection of values for the number of clusters c , weighting exponent m , and constant ν . The algorithm is implemented with $c = 20, 15, 9, 6$ and 5 clusters, the value of m between 1.4 and 1.6 , and the value of ν between 1.0 and 1.5 . The algorithm gave good result with $c = 6, 9; m = 1.5$ and $\nu = 1.0$.

Since paddy transplantation is staggered across the command area, satellite data of any one date does not represent the same growth stage at all locations. In view of this heterogeneity in crop calendar, in order to obtain complete estimate of area under any crop as well as to ensure

better discriminability, satellite data of three dates as mentioned in the previous section are used to reflect the following features.

1. When only semi-dry crops exist (image of 20th February, 1993)
2. When paddy is being transplanted with semi-dry crops already sufficiently grown (image of 14th March, 1993)
3. At the time of maximum ground cover and canopy growth of Paddy (image of 16th April, 1993 shown in Fig. 3).

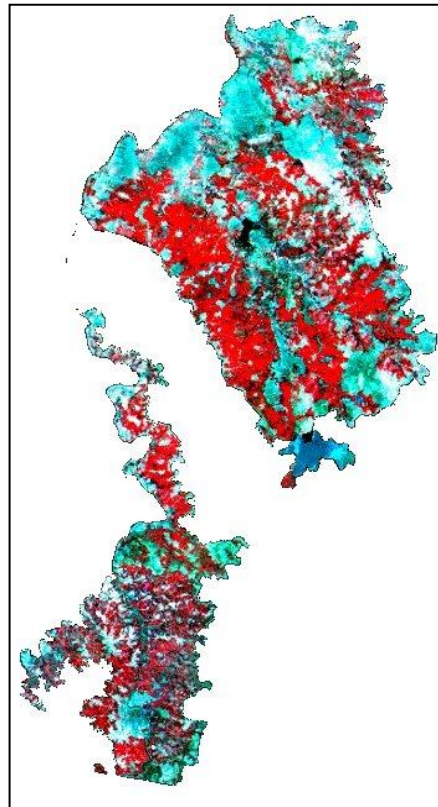


Figure 6 False color composite of study area

(a) Classification of Semi-Dry Crops with Single Date Imagery

Penalized Fuzzy C-Means algorithm was implemented on the 20th February and 14th March multi-spectral images separately to identify semi-dry crops, with $c=5$ and 15 and the results are presented in Table 6.1. As can be noticed $c=5$ is performing better. Areas occupied by semi-dry crops obtained using $c=5$ are given in Table 6.2.

Table 6.1. Semi-dry crop classified using single date imagery with $c = 5$ and 15

Date	c	Available ground truth	Correctly Classified	Misclassified	Accuracy (%)
20 th February, 1993	5	86	84	2	98
	15	86	63	23	73
14 th March, 1993	5	86	82	4	95
	15	86	70	16	81

Table 6.2. Area occupied by Semi-dry crop with $c = 5$

Date	Number of pixels	Resolution (m)	Cropped Area (ha)
20 th Feb, 1993	67,304	72.5	35,376.67
14 th Mar, 1993	75,250	72.5	39,553.28

Using 14th March data, with $c=5$, majority of the Paddy locations were classified into water cluster and some locations to Semi-dry crop because at that time paddy is just transplanted in most of the areas and therefore; water is dominating compared to the crop seedlings. With 16th April data and $c=5$; 42 Paddy locations were correctly classified out of 53 available ground truth locations. Detailed results are given in Laxmi Raju (2003).

(b) Classification of Paddy and Semi-Dry Crops Using Multi-Date Imagery

To discriminate Paddy and Semi-dry crops from other classes, NIR (near infrared) band data of the three dates was used with $c=5$. Only 5 locations were classified into other clusters out of available 53 ground truth locations with 91% accuracy. Similarly for semi-dry crops 95% accuracy is obtained. Classified image is shown in Figure 6.3. Area of Paddy and Semi-dry crops obtained from the NIR band data multi-date imageries are shown in Table 6.3.

Table 6.3. Area of Paddy and Semi-dry crop

Crop	Number of pixels	Resolution (m)	Cropped Area (ha)
Paddy	1,35,405	72.5	71,172
Semi-dry	55,065	72.5	28,943

(c) Classification of Sugarcane Crop

By increasing the number of clusters ($c=15$), it was possible to classify sugarcane crop. However, better results were obtained by using Normalised Difference Vegetation Index (NDVI) images of the three dates, instead of raw data. NDVI is $(NIR - R)/(NIR + R)$ or $(B4 - B3)/(B4 + B3)$ for IRS LISS I. Paddy, Sugarcane and semi-dry crops could be identified from this analysis and the misclassification results (confusion matrix) are presented in the Table 6.4. This table shows the number of pixels from the ground truth, classified into appropriate classes

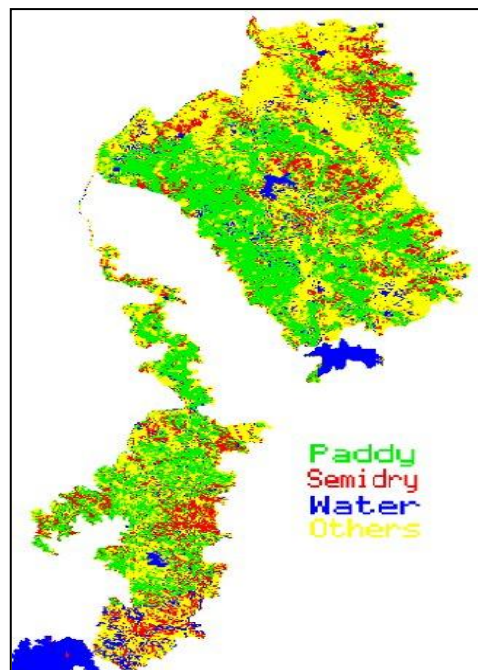


Figure 7 Classified multi-date imagery

Table 4. Misclassification table

	Paddy	Sugarcane	Semidry crop	Other
Paddy	46	3	3	1
Sugarcane	3	14	7	0
Semidry crop	0	12	73	1