CHAPTER 3

PERFORMANCE ANALYSIS OF SUPERVISED ALGORITHMS FOR THE LAND FEATURES CLASSIFICATION

3.1 INTRODUCTION

Crops are the primary and essential necessities for the livelihood and for any growing country of the world. Timely and precise information about diverse crop provide vital and imperative role in the local area crop management during different growing season (Yang et al., 2007). This information is very much expedient for production estimation, timely transportation of crop products and for accurate crop price determination. However, some difficulties may be in separating these crop species due to variations in soil properties, planting dates, fertilization, irrigation, intercropping, pest conditions and tillage practices (Ryerson et al., 1997). Crop classification is associated with the global climate change, agricultural environment and urban development. Therefore, it has attracted widespread interests in the community of geography, ecology, hydrology, Geographical Information System (GIS), remote sensing, environment and onward (Srivastava et al., 2012).

Availability of different satellite imagery, image processing algorithms and advancement in digital image processing enabled the researches and increased the potential to find accurate crop information like as crop type, crop condition and growth of diverse crop in agriculture (Turker and Arikan, 2005; Akbari et al., 2006). The spectral knowledge of diverse crop in satellite imagery is obligatory as a training data for the discrimination of crops in the same or other area (Nidamanuri and Zbell, 2012). High spatial resolution, multi-spectral satellite sensor LISS-IV data in optical and NIR bands emerged as a possible approach for monitoring of crop types and other land

features (Sesha Sai and Narasimha Rao, 2008). Foremost restrictions on agricultural crop type's discrimination by satellite imagery relate to resemblance of reflectance of diverse crop and field-to-field inconsistency of plant reflectance of the same crops. The particular combinations of the crops in a specified region, the array of specific crop phenologies, spatial and spectral variability within fields are also the major limitations (Wheeler and Misra, 1980; Buechel et al., 1989). The LISS-IV sensor data has the potential to capture and remove somewhat these limitations. The usefulness of multispectral satellite imageries have been recognised for the discrimination and mapping of diverse crop by several researchers (WitDe and Clevers, 2004; Conrad et al., 2010; Turker and Ozdarici, 2011).

The crop classification maps are essential for the assessment of amount and type of crops harvested in a certain area and for the management and assessment of agricultural disaster compensation; however, those algorithms have till now to be recognized (Sonobe et al., 2014). Nonparametric algorithms like SVMs, ANN and RF have been found more robust than the conventional statistical algorithms and create no assumptions nearby the statistical nature of the data. These are the good additions to the existing catalogue of image classification algorithms which may allow accurate classification. SVMs algorithm is widely used algorithm in the remote sensing community due to its ability to produce better classification results than the other algorithms even with limited and spectral-mixed training data (Foody and Mathur, 2006; Mountrakis et al., 2011; Shao and Lunetta, 2012). The output of SVMs algorithm depends on the input pixels and pointing out that the training is potentially a significant stage for augmenting classification accuracy (Pouteau and Collin, 2013). However, major anxieties in the design of SVMs classifier such as choice of kernel specific parameters, suitable kernel function, regularization parameter and strategies for

multiclass classification can affect the crop discrimination accuracy results (Huang et al., 2002; Pal 2009; Mountrakis et al., 2011).

ANN algorithm has been well proven as a useful tool in remote sensing and in the other different applications (Mass and Flores, 2008; Kavzoglu, 2009; Gupta et al., 2015; 2016; Mishra and Rai, 2016). The ANN process is a mainly parallel disseminated process which is made from the simple processing units. This has a regular tendency to store experiential information and to make it available in use with unseen data-set (Dastorani et al., 2010; Islam et al., 2012). Further relative studies (e.g. Paola and Schowengerdt, 1995; Mas and Flores, 2008) have demonstrated the superiority of ANNs because they do not assume the training samples to be normally distributed. For example, a back propagation ANN classification process involves a repeated feedforward and back propagation process to minimize the root mean square error (RMSE) between the models predicted and target values. However, few studies have conveyed some problems regarding use of back propagation ANN for crop and other land cover features classification (Foody and Arora, 1997; Kavzoglu and Mather, 2003). ANN algorithms run slowly during the training phase due to involvement of several setting parameters and nature of input data (Arora et al., 2000). The classification accuracy depends on so many factors and may be affected due to variation in the dimensionality of satellite imagery and on training and testing data-sets (Foody and Arora, 1997). RF algorithm is a non–parametric rule based algorithm and it can be trained rapidly. This algorithm is very effective for the classification through complex and non-linear patterns of the landscape. Furthermore, the RF algorithm runs efficiently on large data-sets and does not require normally distributed model training data (Rodriguez-Galiano et al., 2012). SVM needs a several user-defined parameters whereas

RF algorithm requires only two parameters as numbers of trees and number of predictor variables. Classification results by RF were found equally well to SVM (Pal, 2005).

At the advent of a numeral satellites, no insufficiency of data since remotely sensed sources; however, appropriate classification accuracy has remained a big challenge. Despite many research papers published in the field of remote sensing, there is still a lack of comparative studies on different supervised classification algorithms using LISS-IV data. In the presented study, demanding assessment of SVMs with respect to linear, radial basis, polynomial of degree 2, sigmoid function kernels, ANN and RF were endeavoured to perform the crop type and other land features classification and accuracy enhancement by post-processing. The spectral separability along with different bands was analysed for better separation between the crops/non-crops by *TD* and *J-M* distance methods. The crops/non-crops classification accuracy results were investigated statistically by *Z*-test and compared. This study offers significant information regarding the presence of diverse crop in the area dominated by agricultural practices. Such information is vital for the successful management and monitoring of diversity and crop productivity etc.

3.2 DESCRIPTION OF THE STUDY AREA AND MATERIALS

The study area, ground truth information and satellite data used in this chapter is same as described in the chapter 2. Nowadays, it is needed to utilize imagery having high spatial resolution for the better classification. Single-date LISS-IV multispectral imagery can be utilized for the crop classification if the imagery is acquired during the optimum crop discrimination period for a given region. Bearing in mind about the cost of imagery and weather constraints, it is more effective to use single-date imagery for crop discrimination studies if majority of the crops can be covered on a single-date imagery in a region; otherwise multi-date images may be necessary (Yang et al., 2011).

3.3 METHODOLOGY

3.3.1 Image processing of satellite data

Figure 3.1 displays the flowchart of the adopted methodology for the classification of diverse crop and non-crop.

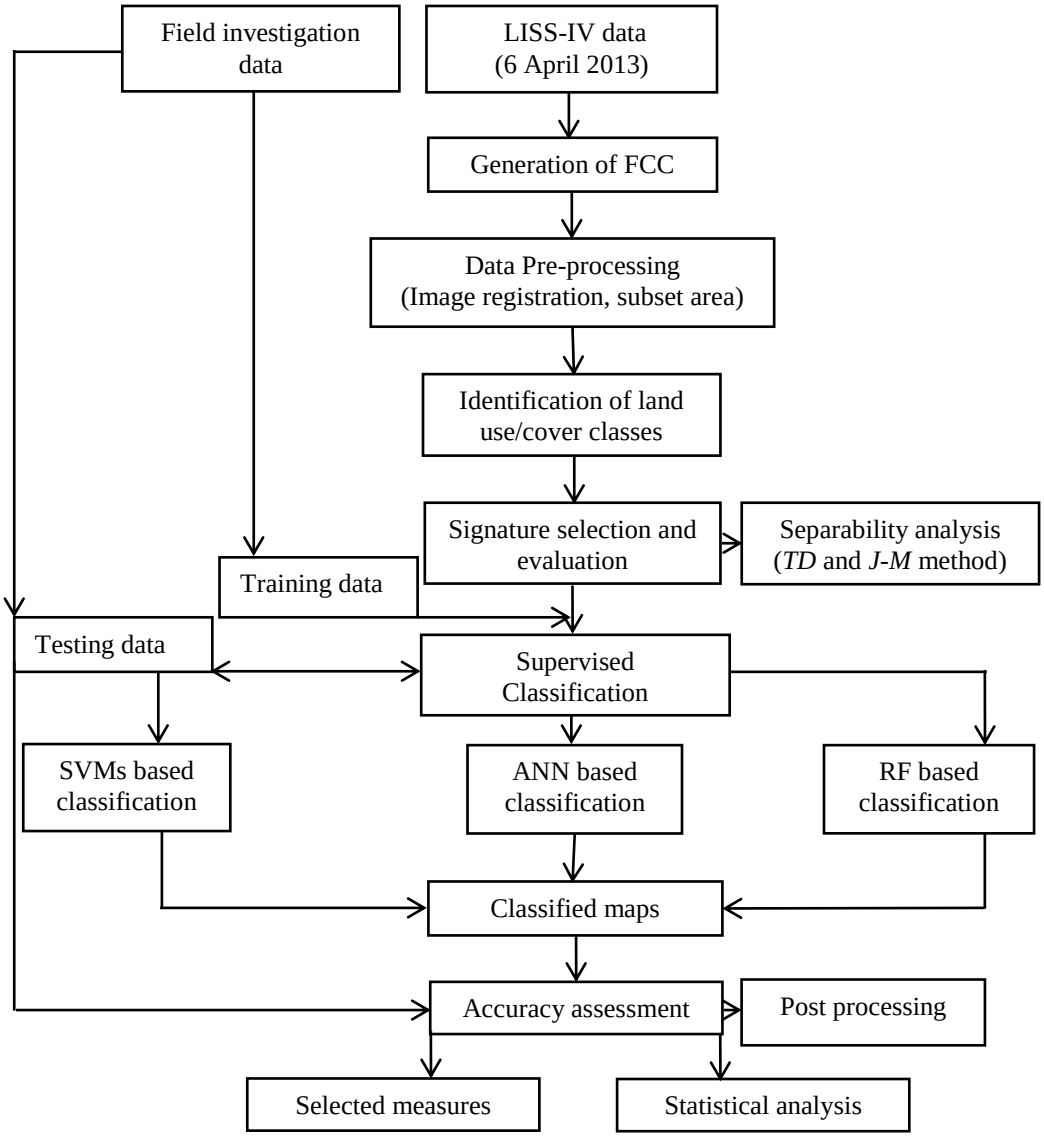

Figure 3.1 Flow-chart of methodology

3.3.2 Image pre-processing

In the remote sensing applications such as image classification, the atmospheric correction was not required in the same calendar date. This was due to atmospheric correction for single date imagery comparable to subtracting a constant from all sample pixels in a spectral band (Song et al., 2001). The layer stacking of three different band

data was done to get the FCC image. The required area was clipped from original image which was found helpful to diminish the dimension of the image file to focus only on the ROI. This procedure not only eliminates the unnecessary data in the file, but also speeds up the processing time, which was much essential for the classification algorithms. The ROI files were generated after the collection of ground samples randomly. The spectral separability before the classification was performed to check the separation between the classes. One ROI file was used to train the classification algorithms and other ROI file was used to test the classifiers. The training/testing pixels used for the discrimination are given in Table 3.1.

Table 3.1 Training and testing pixels used in the classification

Class name	Training pixels	Testing pixels	Class name	Training pixels	Testing pixels
Barley	379	126	Sugarcane	316	108
Wheat	604	200	Other crops	734	247
Lentil	381	125	Water	625	209
Mustard	368	122	Sand	663	220
Pigeon pea	352	111	Built up	462	154
Linseed	337	114	Fallow land	427	142
Corn	491	164	Sparse vegetation	780	269
Pea	385	128	Dense vegetation	739	244

3.3.3 Separability analysis

To define the spectral separability between classes, the *TD* (Swain and Davis, 1978) and *J-M* (Richards, 1999) distance methods were analysed.

3.3.3.1 Transformed divergence method

TD method was used for the class separability analysis between the classes before the image classification. *TD* is a statistical distance measurement between the classes, calculated from means and covariance matrices of each class. The range of *TD* values is from 0.0 to 2.0 and shows how well the chosen training samples are statistically separated. The *TD* values larger than 1.9 show that the classes have good

separability. It was obtained using the equations 3.1 and 3.2 given by Swain and Davis, 1978 as

$$TD_{ij} = 2\left(1 - exp\left(\frac{-D_{ij}}{8}\right)\right) \quad (3.1)$$

D_{ij} = divergence among two signatures and it can be obtained using equation as

$$D_{ij} = \frac{1}{2}\, tr\left(\left(C_i - C_j\right)\left(C_i^{-1} - C_j^{-1}\right)\right) + \frac{1}{2}\, tr\left(\left(C_i^{-1} - C_j^{-1}\right)\left(\mu_i - \mu_j\right)\left(\mu_i - \mu_j\right)^T\right) \quad (3.2)$$

where *i* and *j* are the two classes or signatures being compared, C_i and C_j are the covariance matrices of signatures *i* and *j*, μ_i and μ_j are the mean vectors of signatures *i* and *j*, *tr* indicates about the trace function which computes the sum of the components on the chief diagonal and *T* is the transpose of the matrix used.

3.3.3.2 Jeffries-Matusita distance method

J-M distance allows to indicate how well a selected spectral class pair is statistically separated. The measurement is based on Bhattacharya distance. *J-M* distance for two classes' *a* and *b* was computed by the following equations.

$$JM_{ab} = \sqrt{2(1 - exp(-\alpha))} \quad (3.3)$$

$$\alpha = \frac{1}{8}(\mu_a - \mu_b)^T \left(\frac{C_a + C_b}{2}\right)^{-1} (\mu_a - \mu_b) + \frac{1}{2}\ln\left[\frac{\frac{1}{2}|C_a + C_b|}{\sqrt{|C_a||C_b|}}\right] \quad (3.4)$$

where C_a and C_b are the covariance matrices used for the category *a* and *b*, μ_a and μ_b are the mean values for the category *a* and *b*, and *T* signifies the transpose of a vector. *J-M* distance method delivers a catalogue between 0.0 to 2.0, values > 1.7 demonstrate that the classes are well separated (ITT Industries Inc., 2006). A *J-M* value < 1.0 indicates poor separability between the pair of crops used in the classification.

3.3.4 Image classification

3.3.4.1 Support vector machines based classification

SVMs algorithm is a non-parametric algorithm associated to the statistical learning theory. This algorithm was designed in the late 1970s, nevertheless its fame in remote sensing began to surge about decades ago (Vapnik, 1998; Mountrakis et al., 2011). SVMs utilises a user-defined kernel function to design a set of non-linear verdict boundaries in the original data-set in to linear boundaries of a higher-dimension (Han et al., 2007). This algorithm classifies the sample training data-set into upper dimensional space and finds the best hyperplanes that separate the classes with least classification errors. An optimal hyperplane is firmed using training data-set and its simplification ability is verified using validation data. The best hyperplanes are located using training samples which lie at the boundaries of class distribution in a feature space. The training data-set associated to train the algorithm defines the hyperplane of maximum margin which are called support vectors (Vapnik, 1998; Huang et al., 2002). The remaining training samples can be discarded which are making any involvement to guess hyperplane locations (Brown et al., 2000). An appropriate choice of kernel permits the data to become mostly independent in the feature space in spite of being non-separable in the original input space (Srivastava et al., 2012). The kernel functions, namely linear, polynomial of order 2, radial basis and sigmoidal were used. The gamma parameter (i.e., 0.25) and penalty parameter were set in to their value (i.e., 1000), imposing all the pixels in the training data to unite to a class. The zero identification probability threshold was used for limiting image pixels to acquire exactly one class label and no unclassified pixels remain left (Petropoulos et al., 2011).

3.3.4.2 Artificial neural network based classification

The ANN algorithm is a layered feed forward model in ENVI version 5.1 for the supervised learning. In the satellite image classification, every neuron in the input layer

signifies one input feature as one satellite image band. In case of output layer each neuron resembles to one of the classes to be classified. The back propagation ANN algorithm is one of the utmost normally used forms of neural computing in the remote sensing (Srivastava et al., 2012). ANN is able to process massive amounts of complex and noisy data. The main importance of ANN algorithm is because of its proficiency adaptively to simulate non-linear and difficult patterns with suitable topological structures (Atkinson and Tatnall, 1997). The learning rate parameter was set to 0.01 and momentum value 0.99 for single hidden layer, while stopping criteria was fixed to 0.001. Weights in the ANN were initialized using a uniform distribution, and iteration process was stopped when the RMSE reached at the optimum level. The structure of the ANN is shown in Figure 3.2.

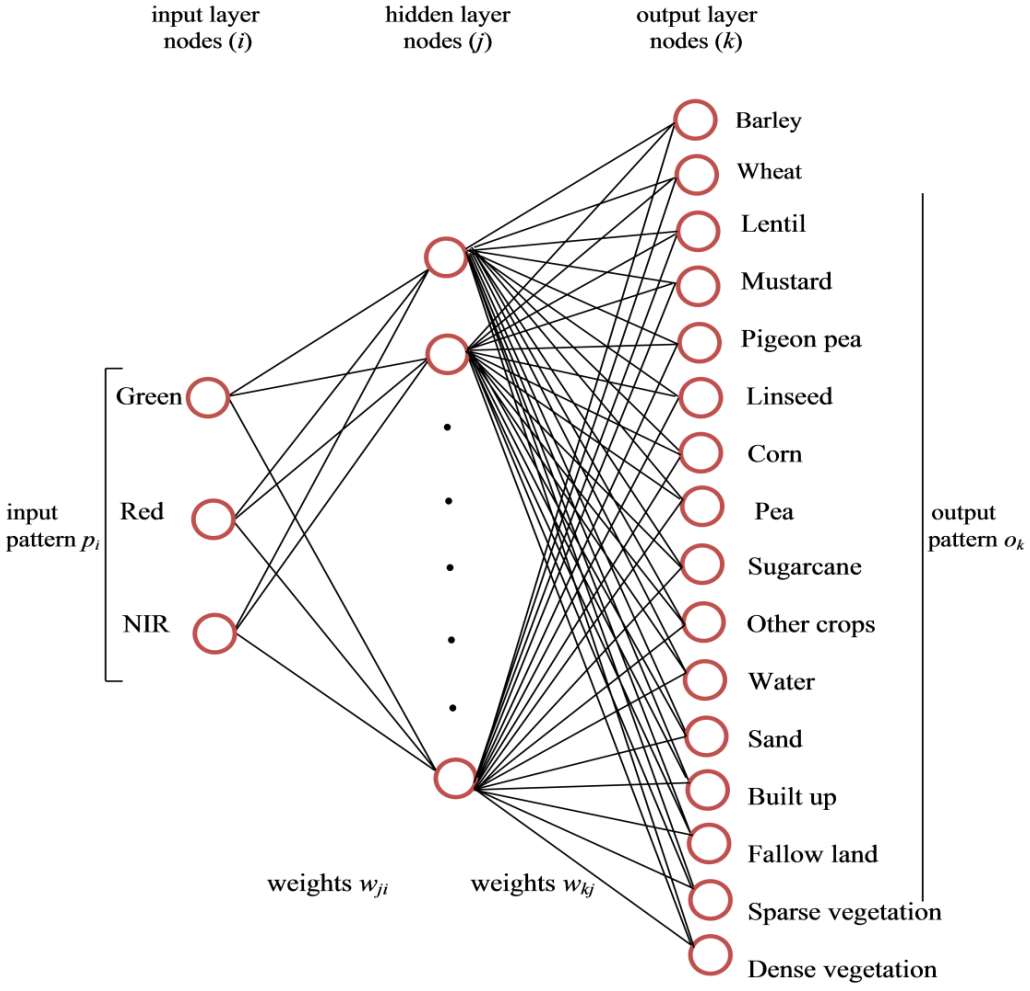

Figure 3.2 The structure of a three- layer ANN

The network has three layers; hidden and output layers contain processing elements at each node. The input layer nodes are simply an interface to the input data and do not do any processing. The input arrays are the features associated in the classification or they are the multispectral trajectories of the training pixels, one band per node (Paola and Schowengerdt, 1995). According to Schowengerdt (2006) within each processing node there is a summation and transformation. At each hidden layer node *j*, the operation performed on the input pattern p_i produces the output h_j,

$$\textit{hidden layer: } S_j = \sum_j w_{ji}\, p_i \tag{3.5}$$

$$h_j = f(S_j) \tag{3.6}$$

which is directed to each output layer node *k*, where the output o_k is calculated as

$$\textit{output layer: } S_k = \sum_j w_{kj}\, h_j \tag{3.7}$$

$$O_k = f(S_k) \tag{3.8}$$

3.3.4.3 Random forest based classification

RF classification algorithm is a robust and an ensemble learning algorithm. The classification analysis using this algorithm is done *via* statistical software R. RF classification algorithm builds many decision trees based on random bootstrapped training sample data-set. Individually, a tree is made of using a different subset from the available training data-set and the nodes in the tree are splited using the best split variable out of a group of randomly selected variables. This approach delivers robustness to over-fitting and can process thousands of dependent/independent variables without deleting variables. To split the nodes user-defined parameters such as number of trees and the number of variables are used. The simplification error always converges on increasing the number of the trees. Therefore, selecting the larger numerals of decision trees is recommended for the RF classification (Breiman, 2001). Some samples do not exist in the training subset which is within another subset named as out-of-bag

(OOB). These remaining OOB elements can also be classified for evaluating the performance of the algorithm by the tree (Rodriguez-Galiano et al., 2012).

3.3.5 Selected measures

3.3.5.1 Marginal rates

A number of asymmetric class measures such as true positive rate (*TPR*) and true negative rate (*TNR*) are reference oriented. All measures lies among 0 to 1. They consider the columns (true classes) of the error matrix and estimated as:

$$TPR_i = n_{TP} / (n_{TP} + n_{FN}) \quad (3.9)$$

$$TNR_i = n_{TN} / (n_{TN} + n_{FP}) \quad (3.10)$$

The correspondent estimation oriented measures such as positive predictive value (*PPV*) and negative predictive value (*NPV*) based on error matrix rows (estimated classes) (Witten and Frank, 2005) are assessed as:

$$PPV_i = n_{TP} / (n_{TP} + n_{FP}) \quad (3.11)$$

$$NPV_i = n_{TN} / (n_{TN} + n_{FN}) \quad (3.12)$$

where n_{TP} (number of true positives) = n_{ii} (number of correctly classified pixels in row *i*); n_{FP} (number of false positives) = n_{i+} (sum of pixels in the error matrix over row *i*) - n_{ii} ; n_{FN} (number of false negatives) = n_{+j} (sum of pixels in the error matrix over column *j*) - n_{ii} ; and n_{TN} (number of true negatives) = *n* (total number of pixels used to test the classification accuracy) - n_{TP} - n_{FP} - n_{FN} (Labatut and Cherifi 2011).

3.3.5.2 *F*-measure

The *F*-measure relates to the harmonic mean of *PPV* and *TPR* (Witten and Frank, 2005). *F*-measure can be evaluated using relation as:

$$F_i = 2\frac{PPV_i \times TPR_i}{PPV_i + TPR_i} = \frac{2n_{TP}}{2n_{TP} + n_{FN} + n_{FP}} \quad (3.13)$$

F measure can be inferred as an extent of overlapping concerning the true and estimated classes.

3.3.5.3 Jaccard Coefficient

Jaccard Coefficient is well known as Jaccard's Coefficient of Community (*JCC*) defined to compare data-sets (Jaccard, 1912). The class specific symmetric measure can be defined as:

$$JCC_i = \frac{n_{TP}}{n_{TP}+n_{FP}+n_{FN}} \quad (3.14)$$

JCC_i is related to the F_i as given in the relation

$$JCC_i = \frac{F_i}{2-F_i} \quad (3.15)$$

3.3.5.4 Classification success Index

The individual classification success index (*ICSI*) is defined for the classification purpose. It is a class specific symmetric measure defined as:

$$ICSI_i = 1\text{-}(1 - PPV_i + 1 - TPR_i) = PPV_i + TPR_i - 1 \quad (3.16)$$

The *CSI* is an overall measure defined by averaging *ICSI* over all the classes (Koukoulas and Blackburn, 2001).

3.4 Statistical significance of classification accuracy by *Z*-test

The *Z*-test is statistical test usually performed to check whether two classification algorithms providing similar classification accuracy. The difference in the crop/non-crop classification accuracy between two algorithms is statistically significant ($p \leq 0.05$) if the *Z* value is greater than 1.96. The value of $Z > |1.96|$ indicates the statistically significant of difference in classification accuracy at the 95% confidence level (Congalton et al., 1983). The statistical significance between two algorithms may be evaluated using equation:

$$Z = \frac{b-c}{N\sigma} \quad (3.17)$$

$$\sigma = \sqrt{\frac{(b+c)-(b-c)^2/N}{N(N-1)}} \quad (3.18)$$

where *N* is the total number of training pixels, *b* represents the number of sample pixels that were correctly classified by the first algorithm but misclassified by the second algorithm. Similarly, *c* represents the sample pixels those were misclassified by first algorithm and found correctly classified by the second algorithm (Kanji, 2006). The number of correctly/wrongly classified sample pixels for two algorithms is presented by Table 3.2.

Table 3.2 Number of correctly and wrongly classified pixels for the two algorithms

Allocation	Classification 2		
Classification 1	Correct	Incorrect	Sum
Correct	*a*	*b*	*a+b*
Incorrect	*c*	*d*	*c+d*
Sum	*a+c*	*b+d*	

3.5 RESULTS AND DISCUSSION

To evaluate the accuracy, actual land cover types for all the fields ~~was~~ were checked on the ground and compared with the pixels or polygons from a classified map developed from satellite data-set. The *OA* and κ of classified images were estimated from error matrices (Congalton and Green, 1999; Jensen, 2005). The *OA* was found by dividing the sum of correctly classified sample pixels to the total reference sample pixels (Lillesand and Kiefer, 1999). The κ is a measurement of how well associated reference data and classified map agree with each other. The strong agreement, moderate agreement and poor agreement occurs if the κ is greater than 0.80, between 0.40-0.80 and less than 0.40 respectively (Jensen, 2005).

Before the discrimination of diverse crop and non-crop, the training samples between the band combinations B2-B3, B3-B4 and B4-B2 were analysed to check the separation among the classes to be classified. The high spectral reflectance in the red and NIR band between classes may be one of the reasons for overall good separation of

the classes using the band combination B3-B4. Representation of training pixels in 2D space between bands 3 and 4 is shown in Figure 3.3.

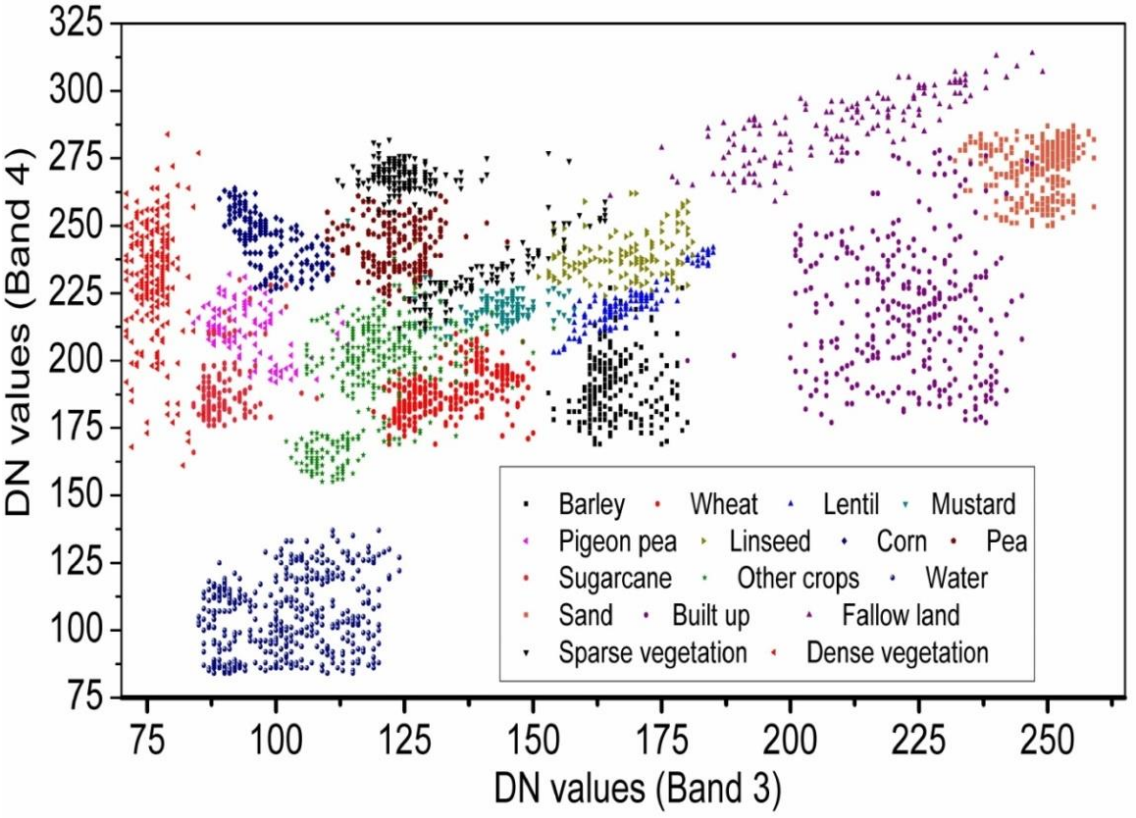

Figure 3.3 Representation of training pixels in 2D space between bands 3 and 4

The two different methods such as *TD* and *J-M* were analyzed and evaluated to check the separability between the classes of crop and non-crop. Almost all the classes were found well separated using *TD* method, however no well separation between all the classes was found using *J-M* method. Although, so many classes provided the same or nearly similar results using *J-M* and *TD* method, but for the some classes, *J-M* method provided less separability between the classes. The *J-M* method provided almost same separability as *TD* method for almost non-crop classes because of its unique spectral response. The separability analysis using *TD* and *J-M* methods are presented in Tables 3.3 and 3.4, respectively.

A comparison between kernels based SVM reveals that the performance of the radial basis function kernel was less affected as compared to linear, polynomial of

degree 2 and sigmoid kernels. The classification accuracy results were also good using RF and ANN algorithms but found lesser in comparison to SVMs except sigmoid kernel. Sixteen-class classification maps generated from the kernel based SVMs algorithm are shown in Figure 3.4.

Table 3.3 Separability analysis between diverse crop and non-crop classes using *TD* method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		1.99	1.97	1.98	2.00	1.98	2.00	1.90	2.00	1.82	2.00	2.00	1.98	2.00	1.97	2.00
2	1.99		1.96	1.95	1.98	1.97	2.00	1.91	2.00	1.73	2.00	2.00	1.98	2.00	1.99	2.00
3	1.97	1.96		1.99	2.00	1.98	2.00	1.96	2.00	1.98	2.00	2.00	2.00	1.98	1.99	2.00
4	1.98	1.95	1.99		1.97	1.71	1.97	1.95	1.99	1.98	2.00	2.00	2.00	2.00	1.97	1.99
5	2.00	1.98	2.00	1.97		1.96	1.98	1.94	1.68	1.77	2.00	2.00	2.00	1.99	1.98	1.47
6	1.98	1.97	1.98	1.71	1.96		1.95	1.96	1.99	1.64	2.00	2.00	2.00	1.88	1.92	2.00
7	2.00	2.00	2.00	1.97	1.98	1.95		1.30	1.93	1.99	2.00	2.00	2.00	2.00	1.84	1.96
8	1.90	1.91	1.96	1.95	1.94	1.96	1.30		1.90	1.98	2.00	2.00	2.00	2.00	1.31	1.99
9	2.00	2.00	2.00	1.99	1.68	1.99	1.93	1.90		1.93	2.00	2.00	2.00	2.00	2.00	1.76
10	1.82	1.73	1.98	1.98	1.77	1.64	1.99	1.98	1.93		2.00	2.00	1.99	1.99	1.92	1.97
11	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00		2.00	2.00	2.00	2.00	2.00
12	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00		1.99	2.00	2.00	2.00
13	1.98	1.98	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.99	2.00	1.99		2.00	2.00	2.00
14	2.00	2.00	1.98	2.00	1.99	1.88	2.00	2.00	2.00	1.99	2.00	2.00	2.00		1.99	2.00
15	1.97	1.99	1.99	1.97	1.98	1.92	1.84	1.31	2.00	1.92	2.00	2.00	2.00	1.99		2.00
16	2.00	2.00	2.00	1.99	1.47	2.00	1.96	1.99	1.76	1.97	2.00	2.00	2.00	2.00	2.00	

Table 3.4 Separability analysis between diverse crop and non-crop classes using *J-M* distance method

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		1.98	1.70	1.88	1.99	1.95	2.00	1.89	2.00	1.71	2.00	2.00	1.97	2.00	1.91	2.00
2	1.98		1.92	1.87	1.97	1.95	2.00	1.85	1.99	1.67	2.00	2.00	1.97	2.00	1.95	2.00
3	1.70	1.92		1.93	1.99	1.87	2.00	1.92	2.00	1.97	2.00	2.00	1.99	1.97	1.82	2.00
4	1.88	1.87	1.93		1.95	1.34	1.95	1.89	1.98	1.92	2.00	2.00	2.00	1.99	1.90	1.97
5	1.99	1.97	1.99	1.95		1.82	1.96	1.57	1.25	1.69	2.00	2.00	2.00	1.98	1.97	1.32
6	1.95	1.95	1.87	1.34	1.82		1.93	1.84	1.98	1.61	2.00	2.00	1.99	1.86	1.71	2.00
7	2.00	2.00	2.00	1.95	1.96	1.93		1.25	1.89	1.98	2.00	2.00	2.00	1.99	1.78	1.82
8	1.89	1.85	1.92	1.89	1.57	1.84	1.25		1.77	1.93	2.00	2.00	2.00	1.99	1.27	1.98
9	2.00	1.99	2.00	1.98	1.25	1.98	1.89	1.77		1.90	2.00	2.00	1.99	2.00	1.99	1.56
10	1.71	1.67	1.97	1.92	1.69	1.61	1.98	1.93	1.90		2.00	2.00	1.96	1.98	1.87	1.95
11	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00		2.00	1.99	2.00	2.00	2.00
12	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00		1.79	2.00	2.00	2.00
13	1.97	1.97	1.99	2.00	2.00	1.99	2.00	2.00	1.99	1.96	1.99	1.79		2.00	1.99	2.00
14	2.00	2.00	1.97	1.99	1.98	1.86	1.99	1.99	2.00	1.98	2.00	2.00	2.00		1.98	2.00
15	1.91	1.95	1.82	1.90	1.97	1.71	1.78	1.27	1.99	1.87	2.00	2.00	1.99	1.98		1.99
16	2.00	2.00	2.00	1.97	1.32	2.00	1.82	1.98	1.56	1.95	2.00	2.00	2.00	2.00	1.99	

1 – Barley, 2 – Wheat, 3 – Lentil, 4 – Mustard, 5 - Pigeon pea, 6 – Linseed, 7 – Corn, 8 – Pea, 9 – Sugarcane, 10 - Other crops, 11 – Water, 12 – Sand, 13 - Built up, 14 - Fallow land, 15 - Sparse vegetation, and 16 - Dense vegetation

Visual comparison of kernel based SVMs, ANN and RF indicated good separations between the classes in the classification maps. Although, it was not visually clear how the crops were separated due to different growing stages and the management conditions of the crop fields. Most of the crop fields had one dominant crop; however some fields contained small presences of other classes.

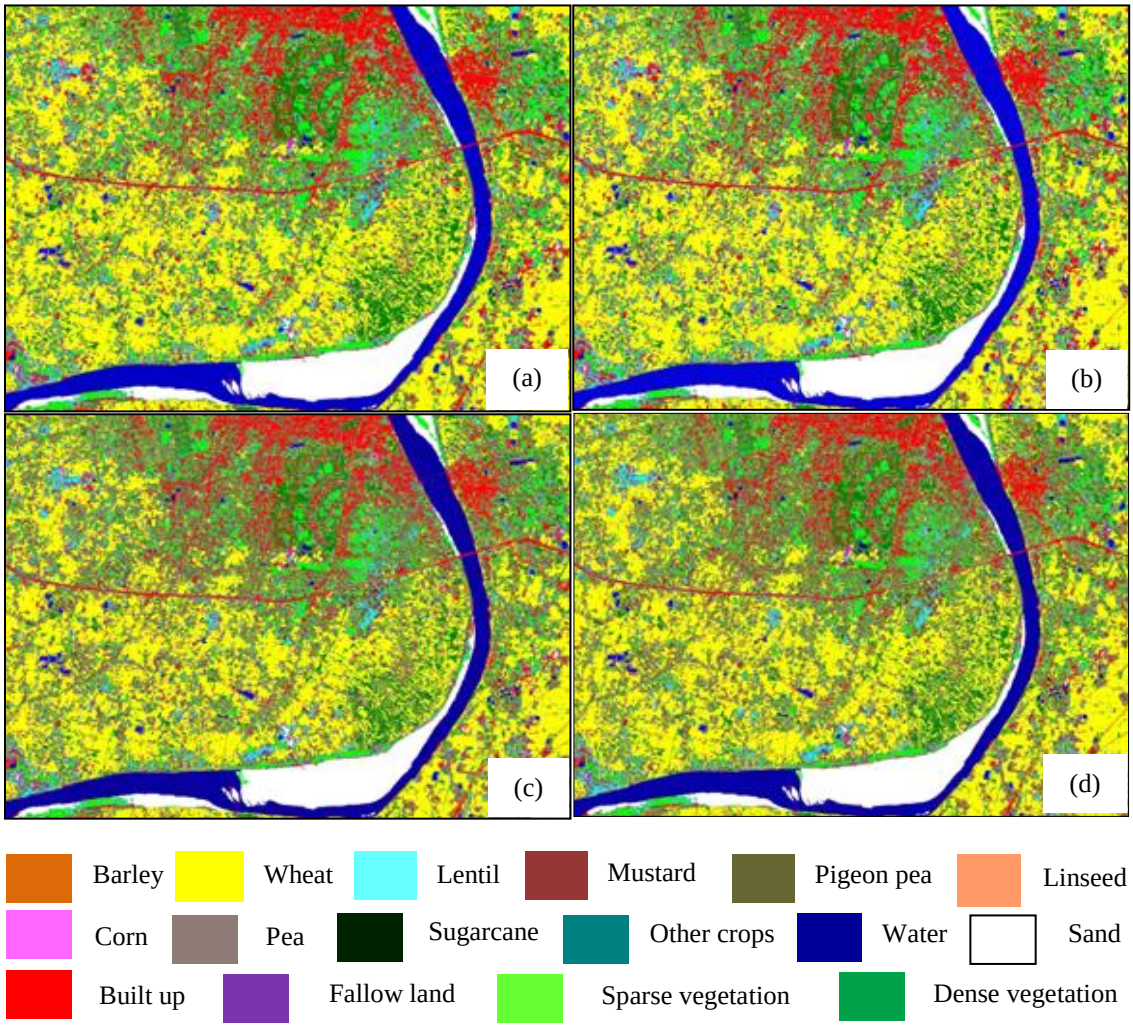

Figure 3.4 (a) SVMs with linear kernel (b) SVMs with polynomial of degree 2 kernel (c) SVMs with radial basis kernel (d) SVMs with sigmoid function kernel based classification

3.5.1 Support vector machines based summary of classification accuracy

Summary of accuracy assessment results for the kernels based SVM maps are presented in Tables 3.5, 3.6, 3.7 and 3.8. The highest *OA* achieved by the radial basis function kernel classified map was 87.89% which specifies the probability of image pixel existence correctly identified approximately 88% in the map. In the case of SVMs

with linear kernel, *TPRi* ranged from 0.5223 for linseed to 0.9446 for wheat in crop classes and from 0.8696 for built up to 1.0000 for water in the non-crop classes. The *PPVi* ranged from 0.6939 for linseed to 0.9692 for wheat in crop classes. It was also ranged from 0.8038 for sparse vegetation to 1.0000 for water in non-crop classes. The same (0.9930) *TPRi* and *PPVi* for fallow land indicate that the fallow land was correctly identified on the ground as well as actually classified on the map.

Table 3.5 Selected measures (SVMs with linear kernel)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.8270	0.9898	0.8014	0.9914	0.8140	0.6863	0.6284
Wheat	0.9446	0.9976	0.9692	0.9956	0.9567	0.9170	0.9138
Lentil	0.8097	0.9910	0.8159	0.9906	0.8128	0.6846	0.6256
Mustard	0.8100	0.9895	0.7838	0.9910	0.7967	0.6621	0.5938
Pigeon pea	0.7054	0.9918	0.7902	0.9872	0.7454	0.5941	0.4956
Linseed	0.5223	0.9899	0.6939	0.9792	0.5960	0.4245	0.2162
Corn	0.8648	0.9897	0.8446	0.9913	0.8546	0.7461	0.7094
Pea	0.7365	0.9914	0.8118	0.9867	0.7723	0.6291	0.5483
Sugarcane	0.8226	0.9891	0.7588	0.9926	0.7894	0.6521	0.5814
Other crops	0.9195	0.9873	0.8800	0.9917	0.8993	0.8171	0.7995
Water	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9909	0.9919	0.9166	0.9992	0.9523	0.9089	0.9075
Built up	0.8696	0.9992	0.9849	0.9922	0.9237	0.8582	0.8545
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.9027	0.9756	0.8038	0.9891	0.8504	0.7397	0.7065
Dense vegetation	0.8820	0.9910	0.9074	0.9881	0.8945	0.8092	0.7894

Table 3.6 Selected measures (SVMs with polynomial kernel)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.7949	0.9906	0.8083	0.9898	0.8015	0.6688	0.6032
Wheat	0.9446	0.9976	0.9692	0.9956	0.9567	0.9171	0.9138
Lentil	0.8255	0.9894	0.7941	0.9914	0.8095	0.6800	0.6196
Mustard	0.8100	0.9891	0.7780	0.9910	0.7937	0.6579	0.5880
Pigeon pea	0.7143	0.9911	0.7770	0.9876	0.7443	0.5928	0.4913
Linseed	0.5219	0.9899	0.6939	0.9792	0.5957	0.4242	0.2158
Corn	0.8463	0.9901	0.8469	0.9901	0.8466	0.7340	0.6932
Pea	0.7439	0.9902	0.7935	0.9871	0.7679	0.6232	0.5374
Sugarcane	0.8129	0.9864	0.7128	0.9921	0.7596	0.6123	0.5257
Other crops	0.9115	0.9873	0.8795	0.9909	0.8952	0.8103	0.7910
Water	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9909	0.9919	0.9165	0.9992	0.9523	0.9088	0.9074
Built up	0.8694	0.9992	0.9850	0.9922	0.9236	0.8580	0.8544
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.9066	0.9756	0.8050	0.9895	0.8528	0.7434	0.7116
Dense vegetation	0.8776	0.9943	0.9391	0.9878	0.9073	0.8303	0.8167

In case of SVMs with polynomial of degree 2, the *OA* and κ were found 87.51% and 0.8658 respectively which were little bit lesser from SVMs with linear kernel but results were found almost similar. The 0.8505 *TPRi* and 0.7222 *PPVi* for sugarcane using SVMs having radial basis kernel function indicate that 85.05% of the sugarcane areas on the ground were correctly identified as sugarcane, but only 72.22% of the areas named sugarcane on the classified map were found actually sugarcane. SVM with sigmoid kernel achieved similar but somewhat lesser accuracy in comparison to all other SVMs kernel functions.

Table 3.7 Selected measures (SVMs with radial basis kernel)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.7559	0.9910	0.8067	0.9879	0.7805	0.6400	0.5626
Wheat	0.9548	0.9976	0.9694	0.9964	0.9620	0.9268	0.9242
Lentil	0.8175	0.9879	0.7687	0.9910	0.7923	0.6561	0.5861
Mustard	0.9091	0.9902	0.8148	0.9957	0.8594	0.7534	0.7239
Pigeon pea	0.6875	0.9930	0.8105	0.9865	0.7440	0.5923	0.4980
Linseed	0.5841	0.9891	0.7021	0.9818	0.6377	0.4681	0.2862
Corn	0.8405	0.9948	0.9133	0.9897	0.8754	0.7784	0.7538
Pea	0.8527	0.9906	0.8209	0.9925	0.8365	0.7190	0.6736
Sugarcane	0.8505	0.9864	0.7222	0.9937	0.7811	0.6408	0.5727
Other crops	0.9113	0.9926	0.9262	0.9910	0.9187	0.8496	0.8375
Water	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9955	0.9923	0.9205	0.9996	0.9565	0.9167	0.9160
Built up	0.8758	0.9996	0.9926	0.9925	0.9306	0.8701	0.8684
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.8806	0.9785	0.8194	0.9866	0.8489	0.7375	0.7000
Dense vegetation	0.8857	0.9930	0.9274	0.9886	0.9061	0.8282	0.8131

The best estimated κ for the SVMs with radial kernel based classification algorithms was 0.8698 which indicates the classification achieved an accuracy that was 87% better than would be expected from random assignment of pixels to classes. The values of all selected measures were found excellent for wheat, sand, other crops, water and fallow land classes, very good for other classes of crop and non-crop and also good results were found for linseed crop. The overall low *TPRi*, *PPVi*, *Fi*, *JCCi* and *ICSIi* for

the linseed crop were mainly because of spectral similarities between lentil, pea, sparse vegetation and other crops in some of the fields.

Table 3.8 Selected measures (SVMs with sigmoid kernel)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.7478	0.9879	0.7539	0.9875	0.7508	0.6011	0.5017
Wheat	0.9400	0.9976	0.9690	0.9952	0.9543	0.9126	0.9090
Lentil	0.7699	0.9867	0.7407	0.9886	0.7550	0.6064	0.5106
Mustard	0.7934	0.9887	0.7678	0.9902	0.7804	0.6399	0.5612
Pigeon pea	0.6874	0.9895	0.7406	0.9864	0.7130	0.5540	0.4280
Linseed	0.5219	0.9867	0.6345	0.9792	0.5727	0.4013	0.1564
Corn	0.8466	0.9893	0.8363	0.9901	0.8414	0.7262	0.6829
Pea	0.7363	0.9902	0.7920	0.9867	0.7631	0.6170	0.5283
Sugarcane	0.7567	0.9884	0.7299	0.9899	0.7431	0.5911	0.4866
Other crops	0.8749	0.9848	0.8543	0.9872	0.8645	0.7613	0.7292
Water	1.000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9909	0.9919	0.9159	0.9992	0.9519	0.9083	0.9068
Built up	0.8693	0.9992	0.9849	0.9922	0.9235	0.8579	0.8542
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.8920	0.9760	0.8050	0.9878	0.8463	0.7335	0.6970
Dense vegetation	0.8817	0.9906	0.9040	0.9881	0.8927	0.8062	0.7857

3.5.2 Artificial neural network based summary of classification accuracy

The ANN algorithm provided nearly similar accuracy results to the kernel based SVMs and RF whereas these results were lowest except SVMs with sigmoid kernel. The classification map produced by the ANN is shown in Figure 3.5. The *OA* and estimated κ using ANN algorithm were found 85.98% and 0.8496 respectively. For the wheat crop, all the algorithms provided TPR_i of nearly 94% indicates that all of the collected validation samples belong to the same class. The wheat crop was found easily in the identification while linseed was the most difficult to differentiate clearly among the crop classes. The unique reflectance value of wheat crop and less mixing of other classes with wheat crop was one of the reasons for identification of this crop easily. The accuracy for linseed crop was found low due to the reflectance values from some of these pixels were similar to the other classes. The *OA* of wheat was found fairly high in comparison to other crop classes; however value 1.0000 of selected measures for water

indicated perfect prediction because of its unique reflectance and no mixing with the other classes in almost all algorithms implemented. The summary of accuracy assessment results are given in Table 3.9.

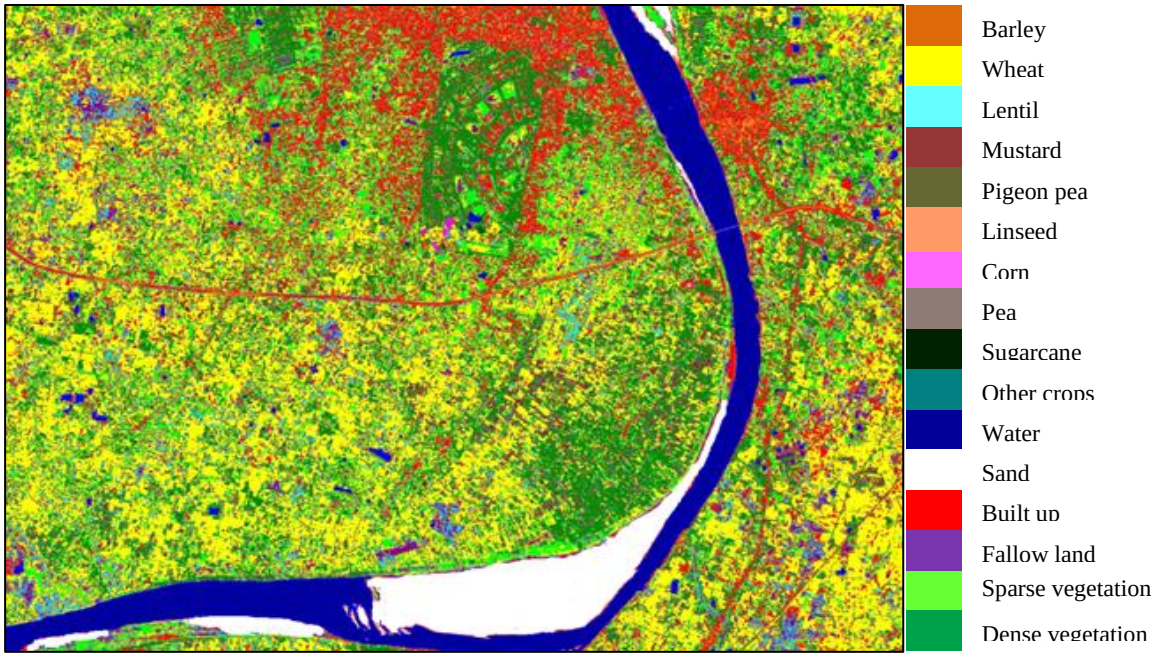

Figure 3.5 ANN based classification

Table 3.9 Selected measures (ANN based classification)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.7640	0.9879	0.7580	0.9883	0.7610	0.6142	0.5220
Wheat	0.9347	0.9980	0.9740	0.9948	0.9540	0.9119	0.9087
Lentil	0.7780	0.9875	0.7540	0.9890	0.7658	0.6205	0.5320
Mustard	0.7932	0.9871	0.7439	0.9902	0.7678	0.6231	0.5371
Pigeon pea	0.7501	0.9895	0.7670	0.9891	0.7585	0.6109	0.5171
Linseed	0.5130	0.9887	0.6666	0.9788	0.5798	0.4083	0.1796
Corn	0.8465	0.9885	0.8264	0.9900	0.8363	0.7187	0.6729
Pea	0.7364	0.9902	0.7920	0.9867	0.7632	0.6171	0.5284
Sugarcane	0.7568	0.9876	0.7170	0.9899	0.7364	0.5827	0.4738
Other crops	0.8789	0.9848	0.8550	0.9876	0.8668	0.7649	0.7339
Water	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9909	0.9919	0.9165	0.9992	0.9522	0.9089	0.9074
Built up	0.8693	0.9992	0.9849	0.9922	0.9235	0.8590	0.8542
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.8989	0.9760	0.8057	0.9887	0.8498	0.7388	0.7046
Dense vegetation	0.8779	0.9947	0.9429	0.9878	0.9092	0.8336	0.8208

The *TNRi* and *NPVi* values were found high for linseed crop whereas the *JCCi* and *ICSIi* measures provided relatively low values. The excellent values of all the measures were found for the wheat, water, sand and fallow land.

3.5.3 Random forest based summary of classification accuracy

The RF classification algorithm based map is presented in Figure 3.6. The *OA* and κ, 86.81% and 0.8582 respectively, were found near to SVMs and ANN algorithms. The TPR_i ranged from 0.4960 for linseed to 0.9400 for wheat crop in crop classes and 0.8693 for built up to 1.0000 for water in non-crop classes. The *PPVi* for crop classes ranged from 0.6914 for linseed to 0.9740 for wheat but it was ranged from 0.8010 for sparse vegetation to 1.0000 for water in non-crop classes.

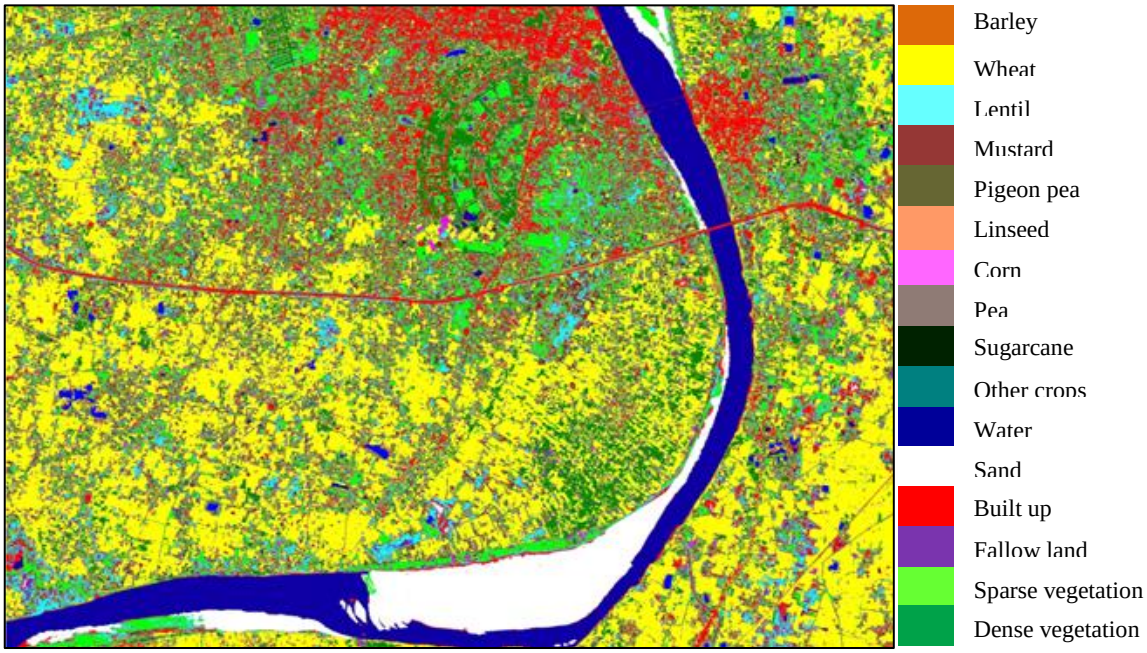

Figure 3.6 RF based classification

The *OA* of wheat crop was highest, because it was the major crop of this season. It may be one of the reasons of high accuracy of wheat crop. The wheat crop has less similarity between sugarcane, pigeon pea, corn and dense vegetation categories. High accuracy results achieved by RF algorithm indicate that there was less mixing among the classes. The corn crop was generally not grown in Varanasi when image acquisition

was done. However, it is grown for the research purpose in the BHU agriculture farm house and some other places in Varanasi. The variability in corn crop fields was originated due to different time of sowing of corn in different fields. It may cause a chance of mixing of corn crop with sparse and dense type of crop classes. The overall high TPR_i and *PPVi* for fallow land indicates less mixing with built up and sand classes. The TPR_i (0.4960) and *PPVi* (0.6914) values for linseed were found lowest using RF in comparison to other classification algorithms. The summary of accuracy assessment results using RF algorithm are given in Table 3.10.

Table 3.10 Selected measures (RF based classification)

Classes	TPR_i	TNR_i	PPV_i	NPV_i	F_i	JCC_i	$ICSI_i$
Barley	0.8190	0.9879	0.7700	0.9910	0.7937	0.6580	0.5890
Wheat	0.9400	0.9980	0.9740	0.9960	0.9567	0.9170	0.9240
Lentil	0.7780	0.9902	0.7970	0.9891	0.7874	0.6493	0.5750
Mustard	0.8019	0.9891	0.7758	0.9906	0.7886	0.6510	0.5777
Pigeon pea	0.6785	0.9918	0.7835	0.9861	0.7272	0.5714	0.4620
Linseed	0.4960	0.9903	0.6914	0.9781	0.5776	0.4061	0.1874
Corn	0.8400	0.9889	0.8305	0.9897	0.8352	0.7171	0.6705
Pea	0.7363	0.9898	0.7849	0.9867	0.7598	0.6127	0.5212
Sugarcane	0.8037	0.9876	0.7291	0.9918	0.7646	0.6189	0.5328
Other crops	0.8989	0.9864	0.8708	0.9897	0.8846	0.7931	0.7697
Water	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sand	0.9864	0.9919	0.9157	0.9988	0.9497	0.9043	0.9021
Built up	0.8693	0.9988	0.9780	0.9922	0.9205	0.8526	0.8473
Fallow land	0.9930	0.9996	0.9930	0.9996	0.9930	0.9861	0.9860
Sparse vegetation	0.9139	0.9747	0.8010	0.9903	0.8537	0.7448	0.7149
Dense vegetation	0.8776	0.9910	0.9069	0.9877	0.8920	0.8051	0.7845

3.5.4 Post-processing summary of classification accuracy

To overcome the problem of mixing, the classified maps were filtered by post processing to eradicate the minor inclusions of other classes within the dominant classes. The different filters such as majority, sieve and clump and the combination of these filters were used to eradicate minor inclusions and for accuracy enhancement. The comparison of the unfiltered and filtered *OA* and κ are given in Table 3.11.

Table 3.11 Classification algorithms *OA*, κ results before and after post processing steps

Classification algorithms		No post processing	Majority filter	Sieve filter	Clump filter	Majority+ sieve	Majority+ clump	Majority+ sieve+clump
SVMs with linear kernel classification	*OA*	87.63%	91.46%	87.59%	88.26%	91.46%	92.10%	92.14%
	κ	0.8670	0.9083	0.8667	0.8740	0.9083	0.9151	0.9155
SVMs with polynomial 2 kernel classification	*OA*	87.51%	91.43%	87.48%	88.15%	91.43%	92.51%	92.58%
	κ	0.8658	0.9079	0.8655	0.8727	0.9079	0.9195	0.9203
SVMs with radial basis kernel classification	*OA*	87.89%	91.76%	87.89%	88.67%	91.76%	92.47%	92.51%
	κ	0.8698	0.9115	0.8698	0.8784	0.9115	0.9191	0.9195
SVMs with sigmoid kernel classification	*OA*	85.76%	89.75%	85.69%	87.48%	89.75%	90.64%	90.83%
	κ	0.8470	0.8899	0.8463	0.8655	0.8899	0.8995	0.9015
ANN classification	*OA*	85.98%	90.36%	85.92%	86.95%	90.36%	91.28%	91.42%
	κ	0.8496	0.8952	0.8489	0.8588	0.8952	0.9048	0.9094
RF classification	*OA*	86.81%	90.76%	86.76%	87.89%	90.76%	91.62%	91.73%
	κ	0.8582	0.9006	0.8577	0.8695	0.9006	0.9094	0.9104

The majority filter increased the *OA* and κ values of all type of classification algorithms. Whereas, the *OA* and κ values were decreased after using sieve filter except for SVMs with radial basis function whose results were found unchanged after the post-processing. Although clump filter provided less *OA* and κ values than the majority filter however, it provided better results than the sieve filter for all classification algorithms. The equal outcomes were found using majority filter and combination of majority with sieve filter. However, the small decrease in the *OA* and κ were found using the combination of majority and sieve filter for the SVMs with sigmoid kernel. The small increment in the *OA* and κ were found using combination of majority, sieve and clump filter in comparison to combination of majority and clump filter. The large mixing of the other crop classes in the dominant crop may be the cause of small decrease in the *OA* and κ using SVMs with sigmoid function kernel. Obviously, image filtering improved not only the *OA* and κ, but also the TPR_i and *PPVi* values for almost all the classes. Nevertheless, an unfiltered classification maps provide well distinction between the classes.

3.5.5 Analyses of statistical significance in the classification accuracy between two algorithms

The almost all classification accuracy combinations provided *Z* values more than 1.96 except SVMs with radial basis kernel *vs.* SVMs with linear kernel ($Z = 1.42$, $p = 0.1556$) and SVMs with polynomial of degree 2 ($Z = 1.66$, $p = 0.0969$) which showed insignificant results. Since *Z* values were less than the value of 1.96, so we do not reject the null hypothesis of no difference. All other combinations were found significantly different with different *Z* and *p* values. The significant increases in the accuracy were found for mustard, linseed, pigeon pea, pea and sugarcane crop. The statistical significance between the algorithms using *Z*-test is presented in Table 3.12.

Table 3.12 Statistical significance in the accuracy between two different algorithms by Z-test

Classification1	Classification 2	Z-test	*p* value
SVMs with radial basis kernel	SVMs with linear kernel	1.42	= 0.1556
SVMs with radial basis kernel	SVMs with polynomial of degree 2	1.66	= 0.0969
SVMs with radial basis kernel	SVMs with sigmoid kernel	4.08	< 0.0001
SVMs with radial basis kernel	ANN classification	4.02	= 0.0001
SVMs with radial basis kernel	RF classification	3.86	= 0.0001

The combinations using SVMs with radial basis kernel *vs.* SVMs with linear kernel and SVMs with polynomial of degree 2 were more accurate than the other combinations. The *Z* values were found less than 1.96 at the 95% confidence level for these two combinations.

3.6 CONCLUSION

The kernel based SVMs, ANN and RF provided nearly similar results in the present study. The unfiltered classified maps provided fairly good crop classification. However, results were much improved after filtering the classified maps which eradicates the inclusions of other classes within the dominant class. The results were significant except SVMs having radial basis function vs. SVMs having linear kernel and polynomial kernel of degree 2 by *Z*-test. The LISS-IV sensor was found much useful for the discrimination of diverse crop like corn, lentil, linseed, barley, mustard, pigeon pea, wheat, sugarcane, pea and other crops and also for non-crop using different supervised

classification algorithms. The high spatial resolution of LISS-IV sensor makes it easier to delineate boundaries among various crop fields. It also enabled improvement for crop discrimination by avoiding overlapping of land covers. Forthcoming investigation should be more focused on accuracy enhancement of linseed crop and some other less accuracy classes with involvement of some more recent algorithms. The approaches and algorithms presented for crop discrimination and mapping can be useful for different crop grown in the other regions.