Comparison between Possibilistic c-Means (PCM) and Artificial Neural Network (ANN) Classification Algorithms in Land use/ Land cover Classification

Ganchimeg Ganbold*, Stanley Chasia**

ARTICLE INFO

ABSTRACT

There are several statistical classification algorithms available for land use/land cover classification. However, each has a certain bias or compromise. Some methods like the parallel piped approach in supervised classification, cannot classify continuous regions within a feature. On the other hand, while unsupervised classification method takes maximum advantage of spectral variability in an image, the maximally separable clusters in spectral space may not do much for our perception of important classes in a given study area. In this research, the output of an ANN algorithm was compared with the Possibilistic c-Means an improvement of the fuzzy c-Means on both moderate resolutions Landsat8 and a high resolution Formosat 2 images. The Formosat 2 image comes with an 8m spectral resolution on the multispectral data. This multispectral image data was resampled to 10m in order to maintain a uniform ratio of 1:3 against Landsat 8 image. Six classes were chosen for analysis including: Dense forest, eucalyptus, water, grassland, wheat and riverine sand. Using a standard false color composite (FCC), the six features reflected differently in the infrared region with wheat producing the brightest pixel values. Signature collection per class was therefore easily obtained for all classifications. The output of both ANN and FCM, were analyzed separately for accuracy and an error matrix generated to assess the quality and accuracy of the classification algorithms. When you compare the results of the two methods on a per-class-basis, ANN had a crisper output compared to PCM which yielded clusters with pixels especially on the moderate resolution Landsat 8 imagery.

1. Introduction

Economic globalization combined with the looming global land scarcity increases the complexity of future pathways of land use change. Predictions of the expected land use impact of national policies have become more uncertain (Lambin & Meyfroidt, 2011). In a more interconnected world,

 * Senior Lecturer, E-Open Institute, Mongolian University of Science and Technology, Mongolia (ganaa_mzb@yahoo.com)
** Assistant Lecturer, Department of Geosciences and the Environment, Technical University of Kenya, Kenya (schasia@gmail.com)

agricultural intensification may cause more rather than less cropland expansion. Land use regulations to protect natural ecosystems may merely displace land use elsewhere by increasing imports (Lambin & Meyfroidt, 2011). Knowledge about land use and land cover has become increasingly important as nations plan to overcome the problems of haphazard, uncontrolled development, deteriorating environmental quality, loss of prime agricultural lands, destruction of important wetlands, and loss of fish and wildlife habit at (Anderson et al., 1976). In addition, satellite-derived data is now capable of linking land cover with socioeconomic and environmental indicators (Grekousis, Mountrakis, & Kavouras, 2016). This information offers planning agencies with accurate data that would inform policy guidelines and ensure sustainable development. Land use data are needed in the analysis of environmental processes and problems that must be understood if living conditions and standards are to be improved or maintained at current levels (Anderson et al., 1976). Remotely sensed images are attractive sources for extracting land cover information, where an image classification algorithm is employed to retrieve land cover information (Debojit, Arora Manoj, & Balasubramanian, 2011). They represent an important, cheap and no time consuming font of data (Follador et al., 2008). Automated land cover/land use change detection from multi-temporal satellite data is one of the most important challenges facing the remote sensing community (Ndehedehe et al., 2013). In the past few years, satellite image classification to produce land use or land cover maps has shifted from finding the right data to finding a method able to cope with the plethora of available data (Stathakis & Vasilakos, 2006). Characteristics related to spatial resolution, overall accuracy, time of data acquisition, sensor used, classification scheme and method, support for land cover change detection, download location, and key corresponding references are provided (Grekousis, Mountrakis, & Kavouras, 2015). Several classification algorithms are used in remote sensing. Many studies have demonstrated the effectiveness of Artificial Neural Networks (ANN) in remote sensing classi- fication (Pratola et al., 2011). In this research, the output of an ANN algorithm was compared with the Possibilistic c-Means, an improvement of the fuzzy c-Means, on both moderate resolutions Landsat 8 and a high resolution Formosat 2 images.

Objectives: To assess and compare the robustness of both the possibilistic c-Means and the Artificial Neural Network (ANN) classification algorithms in extracting landuse/landcover classes in the study area.

- To assess the effectiveness of Possibilistic c-Means and ANN classifiers in extracting landuse/landcover classes
- To extract landuse/land cover classes using ANN and Possibilistic c-Means classifier for Formosat 2 and Landsat 8 images
- To compare and evaluate sub-pixel accuracy information using Possibilistic c-Means and ANN algorithms

Traditional scientific remote sensing classification techniques rely heavily on statistical algorithms to identify and classify pixels on a given set of classes. Some methods like supervised classification rely on human intervention and are therefore biased and/or subjective in nature. There are also errors that occur especially in cases where local knowledge of the study area is lacking. Also, satellite images can enhance the quality by using other methods of image processing (Ganchimeg, 2015).

2. Methods

2.1 Artificial Neural Network

The ability of the human brain and eye to recognize features can be modeled using the Artificial Neural Network (ANN). The neurons are trained to recognize individual pixels and cluster them correctly in their respective classes. This is a very complex process especially when a multi-layer hierarchical system is used. A well trained ANN however, is capable of performing classification better than a human being since it is devoid of subjectivity and human error. ANN can be stated as

$$net_i = \sum_j w_{ij} * O_j + b_i \quad (1)$$
$$O_i = f(net_i)$$

Every connection in the network has a numerical value attached to it called weight (w_{ij}). A unit i computes a net input (net_i) from the outputs o_j of other units. A bias (b_i) which is a numerical value is normally added to the net input. A function f is applied to this value yielding o_i as the output of the unit.

Artificial Neural Networks (ANNs) technology is an alternative to constructing a computer-based simulation system for land classification (Huang & Lippmann, 1987; Hepner et al., 1990; Gong, Pu, & Chen, 1996). Since the beginning of the 1990s, ANNs, also known as neural networks, have been applied to the analysis of remote sensing images with promising results (Atkinson & Tatnall, 1997). ANN is an empirical modeling tool that has an ability to identify underlying highly complex relationship from input output data only (Aqil et al., 2006). ANNs are suitable for analysis of virtually every data type, regardless of their statistical properties (Xie, Sha, & Yu, 2008). ANN have the advantage of a high computation rate due to its massive parallelism as a result of the dense arrangement of interconnections *(weights)* and simple processors (neurons), which permits real-time processing of very large data sets (Mather & Tso, 2009). The spatial tracing and location analysis of emergency incidents is achieved through the utilization of an ANN. More specifically, the ANN provides the basis for a spatiotemporal clustering of demand, definition of the relevant centers, formulation of possible future states of the system and finally, definition of locational strategies for the improvement of the provided services (Photis & Grekousis, 2012). ANN have been used to solve complex problems in location analysis (Photis & Grekousis, 2012), and emergency analysis (Grekousis & Photis, 2014; Hsu & Li, 2010). Furthermore, ANN have also been successfully applied in the classification of remotely sensed Images, particularly in land-use change (Gong, Thill, & Liu, 2015), forest-fire classification, geological mapping and urban area classification (Paola & Schowengerdt, 1995; Mather, 1999).

2.2 Multilayer Perceptron (MLP)

A Multilayer Perceptron (MLP) is a network of simple neurons called perceptron. The basic concept of a single perceptron was introduced by Rosenblatt in 1958. MLP is an important class of ANN; in fact it is the most used in all applied fields. It is a feed forward ANN model based on supervised training. An MLP consists of a set of input units *(the input layer)*, one or more

sets of computation nodes *(the hidden layers)*, and one set of computation/output nodes *(the output layer)*. Connections are always made forward, on a layer-by-layer basis (Mas & Flores, 2008). MLP networks are architectures in which each node receives inputs from previous layers and information flows in one direction to the output layer (Pratola et al., 2011). The number of nodes in the intermediate layer*(s)* defines both the complexity and the power of a neural network model to describe underlying relationships and structures inherent in a training data set (Kavzoglu, 2009) and what more nodes in such layers may be required for classification of more complex, grainy satellite images (Jarvis & Stuart, 1996). Training or learning is done by updating the connection weights in an iterative manner based on certain algorithms; the common one being used is the Back propagation algorithm (Gonzalez & Richard, 2016).

2.3 Fuzzy algorithms

Fuzzy C-Means (FCM) was proposed by Dunn in 1973 and was modified by Bezdek in 1981. It is one of the most popular fuzzy clustering techniques with the approach that the data points have their membership values with the cluster centers that will be iteratively updated (Chattopadhyay, Pratihar, & De Sarkar, 2011). FCM clustering involves two major steps: the calculation of cluster centers and the assignment of points to these centers using a form of Euclidian distance such that the process is continuously repeated until the cluster centers stabilize (Thomas & Nashipudimath, 2012). The FCM algorithm provides a method of clustering that enables a data item to belong to two or more clusters and this scheme of method is frequently used in pattern recognition applications (Velmurugan, 2012). Although the FCM was originally proposed as a clustering *(unsupervised)* technique, the algorithm may be modified so that the classification is based on class canters provided by the analyst from training samples and so for use as a supervised classifier (Foody, 1995a). Indeed, the FCM has been commonly used in a supervised mode to derive sub-pixel scale thematic information from remotely sensed data (e.g. Foody, 1995b; Atkinson et al., 1997; Bastin, 1997; Lucas et al., 2002). Although accurate estimates of sub pixel class composition have been derived with the FCM it is apparent that accuracy is a function of the value of *m* used in the analysis (Foody, 1995b) and thus this value should be carefully selected for the application in-hand.

The possibilistic C-Means (PCM) is the possibilistic counterpart of the FCM. The main difference between the PCM and the more widely used FCM is the removal of the constraint for the memberships to sum to one or each pixel in the PCM (Krishnapuram & Keller, 1993; Foody & Cutler, 2006). The main attraction of the PCM for the derivation of sub pixel scale thematic information is that, like the typicality probability, the membership values derived are measures of the absolute strength of class membership (De Jong & Van der Meer, 2007). Consequently, the memberships derived from the PCM are not affected by the presence of untrained classes. Although more accurate predictions of sub-pixel class composition may be derived from the FCM than PCM, the presence of untrained classes can markedly degrade the accuracy of sub-pixel estimates from the FCM while not affecting those from the PCM (De Jong & Van der Meer, 2007). Thus, in situations when the analyst may believe that the set of classes contained within the imaged area has not been defined exhaustively

it may be preferable to use the PCM rather than FCM for the derivation of sub-pixel scale thematic information (Foody, 2001). PCM enables clustering of noisy data samples i.e. data sets with presence of outliers or noisy points (Suganya & Shanthi, 2012). Each cluster therefore is independent of the other clusters. The objective function corresponding to cluster *i* can be formulated as

$$J_i(\beta_i, U_i, X) = \sum_{j=1}^{N} (u_{ij})^m d^2(x_j, \beta_i) + \eta_i \sum_{j=1}^{N} (1 - u_{ij})^m \quad (2)$$

In (2), β represents the prototype associated with cluster *i*, U_i represents the *i* -th row of the membership matrix *U*, and η is the resolution parameter.

The membership values in PCM can be interpreted as the degree of belongingness, compatibility or typicality (Chawla, 2010). PCM is extremely sensitive to good initialization (Grover, 2014). Some of the key differences between PCM and FCM are outlined below:

- FCM is considered primarily a partitioning algorithm (Krishnapuram & Keller, 1996). It will find a fuzzy C-partition of a given data set, irrespective of how many clusters are present in the data set.
- In PCM, each component generated by the PCM corresponds to a dense region in the data set. The prototypes are automatically attracted to dense regions in feature space as iterations proceed.
- Fuzzy clustering algorithms such the FCM have constraints in membership which causes them to generate memberships that can be interpreted as degree of sharing as opposed to degrees of typicality (Krishnapuram & Keller, 1996). The downside to this is effect is that it gives rise to poor performance in the presence of noise and outliers.
- In the possibilistic approach, membership value of a point in a cluster or class represents the possibility of the point belonging to the class.
- In the case of untrained classes during supervised classification, FCM is less efficient than PCM because it gives the relative membership value as opposed to the absolute.

3. Study area

The study area is the Haridwar District, found in the State of Uttarakhand (Fig 1) in the northern part of India. It covers a geographical area of about 2,360km². It shares it's boundaries with Dehradun to the north, Pauri-Garhwal to the east, while, the west and south areas are bounded by the Uttar Pradesh District.

The central latitude and longitude of the district are 29.956 N and 78.170 E respectively. The river Ganges flows through the district in a series of channels separated from each other called *aits*, most of which are well wooded. Other minor seasonal streams are *RanipurRao, PathriRao, RawiiRao, HarnauiRao, and BegamNadi* among others. A large part of the district is forested, with a population of 1,927,029 inhabitants, has summer temperature ranges of 35 °C - 42 °C and 6 °C - 16.6 °C during winter. The district receives monsoon rains, mostly during the summer. The study area was selected for this project work using Formosat 2 data is Haridwar, Uttarakhand and is shown in Figure 1.
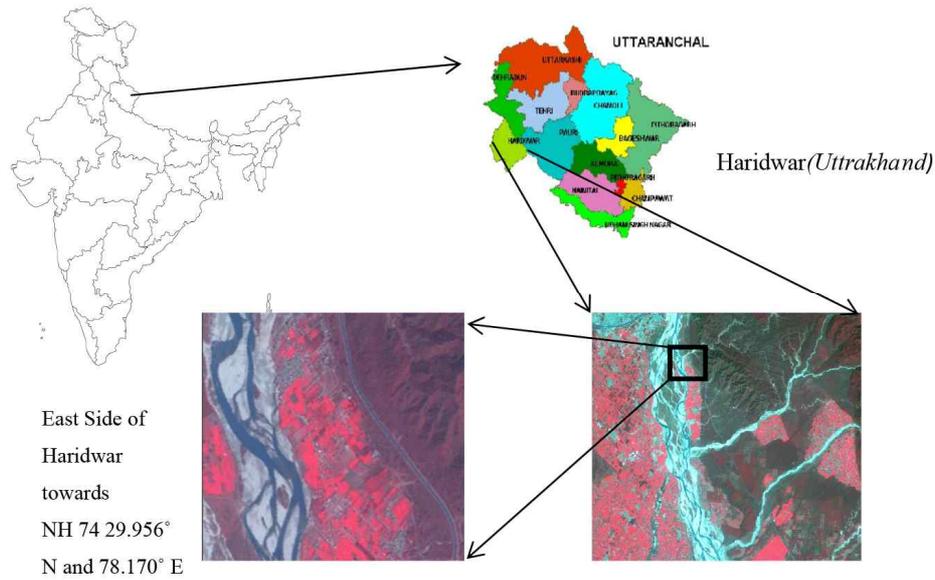
Fig. 1. Study area selected, Formosat 2 data

The landuse/landcover classes in the study area include: eucalyptus trees, fallow land, forested land, water and wheat.
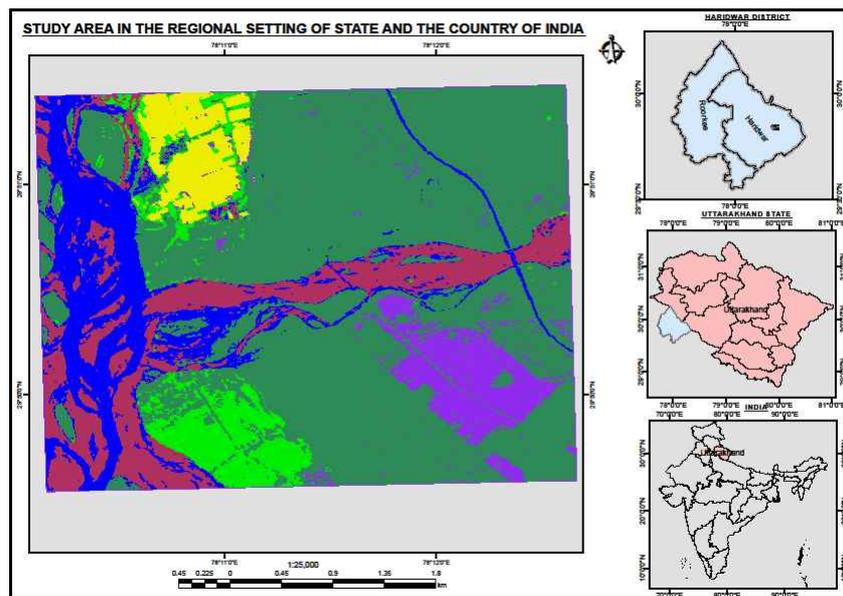
Fig. 2. Study area Haridwar District in Uttarakhand

Datasets: The data used in this research project is Landsat 8 *(moderate resolution)* and the Formosat 2 data *(high resolution)* for comparison purposes. Landsat 8 is an Earth observation satellite, launched

on February 11, 2013. It is the eighth in a series of satellite in the Landsat program and the seventh to reach orbit successfully, providing moderate-resolution imagery, from 15 meters to 100 meters, of Earth's land surface and Polar Regions. Figure 3, shows the imagery of the study area in False Color Composite (FCC), dated 12th February, 2015.
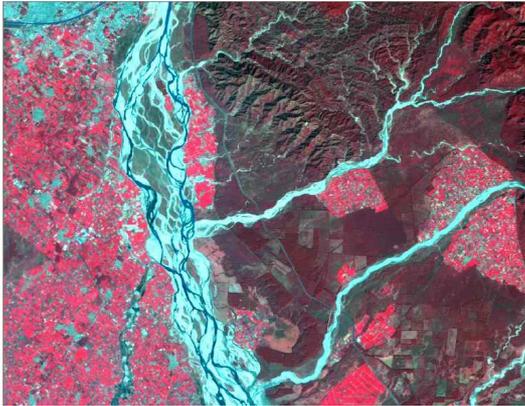
Fig. 3. Image of Haridwar and surroundingstaken from Landsat 8
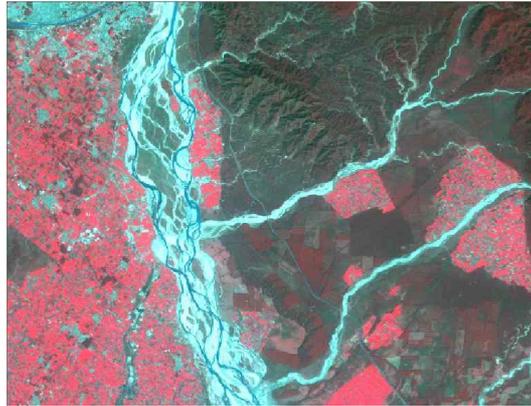
Fig. 4. Image of Haridwar and surroundingstaken from Formosat 2

Landsat 8 operates in the visible, near-infrared, short wave infrared and thermal infrared spectrums. Following are some of its key features:

- *Output format*: GeoTIFF
- *Pixel size*: 15 meters/30 meters/100 meters (panchromatic / multispectral / thermal)
- *Map projection*: UTM (Polar Stereographic for Antarctica)
- *Datum*: WGS 84
- *Resampling*: Cubic convolution.

The sensor spectral bands specifications have been listed in Table 1.

Table 1. Spectral Bands of Landsat8

Spectral Band	Wavelength	Resolution
Band 1 - Coastal / Aerosol	0.433 - 0.453 μm	30 m
Band 2 - Blue	0.450 - 0.515 μm	30 m
Band 3 - Green	0.525 - 0.600 μm	30 m
Band 4 - Red	0.630 - 0.680 μm	30 m
Band 5 - Near Infrared	0.845 - 0.885 μm	30 m
Band 6 - Short Wavelength Infrared	1.560 - 1.660 μm	30 m
Band 7 - Short Wavelength Infrared	2.100 - 2.300 μm	30 m
Band 8 - Panchromatic	0.500 - 0.680 μm	15 m
Band 9 - Cirrus	1.360 - 1.390 μm	30 m

For the comparative study of the different norms using FCM classification, the data set of Formosat 2 has also been used for the same area. This was the first remote sensing satellite developed by National Space Organization (ISPO). Formosat 2 satellite carries both *"remote sensing"* and *"scientific observation"* tasks in its mission. It supports monitoring and detecting land change for any specific regions for various industries and mapping applications. Figure 4, shows the imagery of the study area in False Color Composite (FCC), dated February 21, 2015.

Formosat 2's ability to acquire repeat imagery of an area of interest every day and with the same viewing parameters guarantees a timely flow of compatible data, letting one analyze and compare imagery acquired at different dates with no need of additional processing. The satellite captures panchromatic and multispectral data simultaneously with 2m and 8m spatial resolution respectively. The sensor footprint is 24*24 km and is designed in such a way to revisit the same point on the globe every day in the same viewing conditions. The sensor spectral bands specifications have been listed in Table 2.

Table 2. Formosat 2 Sensor Specifications

Band	Wavelength (micrometers)	Spatial Resolution (meters)
Band 1 - Blue	0.45 - 0.52	8
Band 2 - Green	0.52 - 0.60	8
Band 3 - Red	0.63 - 0.69	8
Band 4 - Near Infrared (NIR)	0.76 - 0.90	8
P - Panchromatic	0.45 - 0.90	2

The following are the six (6) land use classes used in the study (Table 3):

Table 3. Landuse classes

S. No.	Land Use / Land Cover Categories
1	Grassland
2	Wheat
3	Eucalyptus Trees
4	Riverine sand
5	Water
6	Dense forest

4. Methodology

The analysis process began by geometrically correcting both Landsat 8 image and Formosat 2, to a common datum using a 1: 50,000 topographical map of the Dehradun area. This output was then corrected for radiometric errors due to the difference in pixel intensity values attributed to different illumination angles and acquisition date and time.

A subset was performed on the two corrected images to extract only the extent of the study area. This reduces redundancy in data processing while saving time and storage space. The multispectral bands of Formosat 2 data were resampled from a spectral resolution of 8 meters down to 10 meters. This was aimed at maintaining a standard ratio of 3:1 *vis a vis* Landsat 8 *(30m)* data for easy comparison.

Six *(6)* signature classes were developed for the six target features for classification: *Eucalyptus, wheat, water, forest, riverine sand and grass*.

The data used for this study include Landsat 8 *(moderate resolution)* and the Formosat 2 data *(high resolution)* with both panchromatic and multispectral bands at 2m and 8m spatial resolution respectively covering a footprint of 24*24 km. The two data sets were selected on the basis of their availability and also to draw comparisons between the different classes under classification.

Classification signatures were developed for the 6 classes by collecting several samples pixels per class.
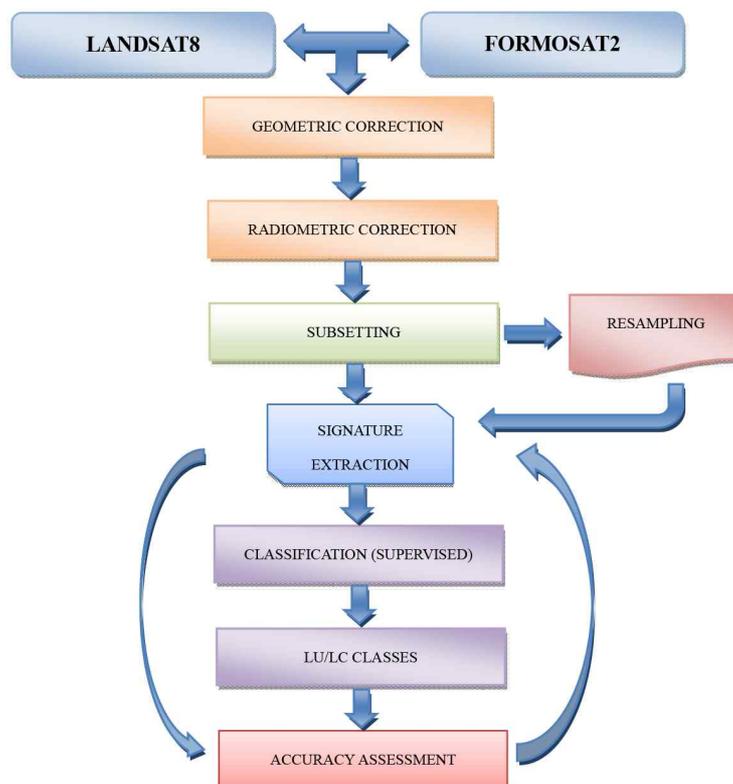
Fig. 5. Data analysis flow

The signatures were evaluated for accuracy using the feature space to image masking and the signature methods. Signatures with overlapping pixels i.e. pixels values that appear similar in a feature space were edited or recreated altogether.
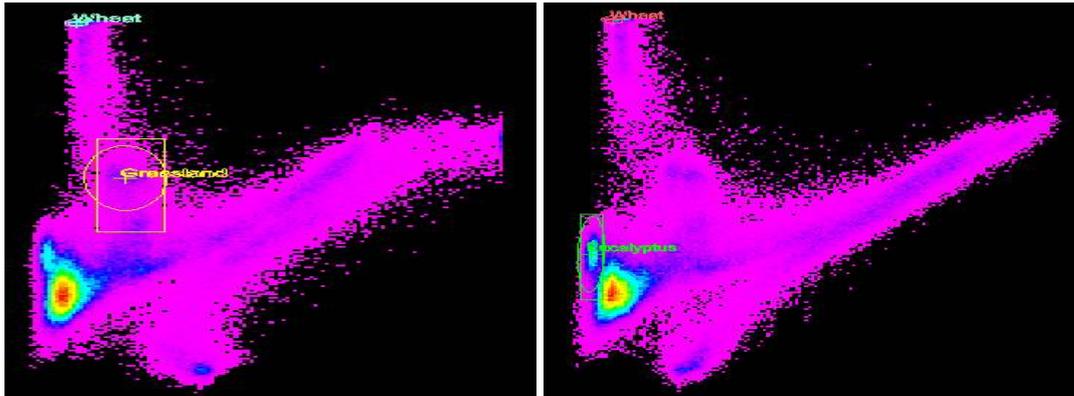
Fig. 6. Signature seperability

The supervised classification was then conducted on the images to extract the different land use/land cover classes based on the edited signature. The output was then evaluated for accuracy using the Cohen's Kappa coefficient which measures the agreement between two raters each classifying *n* items into mutually exclusive categories. This is represented with the equation (3) below.

$$k = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \qquad (3)$$

Where:

P_o - Overall accuracy from the confusion matrix

P_e - Hypothetical probability of chance agreement

If the results are not satisfactory, the signatures are recollected afresh and the entire classification process is repeated. Using artificial neural network algorithm on both Landsat 8 and Formosat 2 data, 6 files membership images corresponding to the classified classes and one hard classified image were produced. The membership files displayed the feature in a panchromatic mode with brighter cells depicting the presence and concentration of the feature. Resultant fraction images after classifying Landsat 8 with Artificial Neural Network (ANN).

5. Results and Discussion

Neural networks are information processing systems which consist of a large number of very simple yet highly interconnected processing elements called units. For training and testing of the classification accuracy of ANN, the research used a resampled subset scene of Landsat 8 data and Formosat 2. The aim of using the neural network was to distinguish land-use classes with more or less similar pixel intensity especially on the moderate resolution Landsat data. Two thematic maps (Fig 10) were prepared from both Landsat and Formosat data by visual classification using

field data. From the two scenes, slightly over half of the pixels per land-use class were chosen for the training of the neural network (Table 6). This research used 1 hidden layer network with a root mean square (RMS) exit criteria of 0.1000. The numbers of output nodes in the network were six corresponding to the land-use classes: wheat, riverine sand, water, eucalyptus, dense forest and grassland.

The network was trained with the back-propagation algorithm using a cross-validation approach where the training data set was subdivided into validation and training subsets. The network learned from the training subset and stopped at several points during the learning process. At each stopping point, the network was used to classify the samples contained in the validation subset. This training continued until the classification error of the validation subset began to rise.

For supervised PCM classification, the number of training sites varied per class and image type. More samples were generally picked on the high resolution Formosat image compared to the moderate resolution Landsat 8 data, where fewer training sites per class were picked. This is attributed to the high spectral variability in Formosat which has an effect of fuzziness. The membership values for each pixel in the different land-use classes ranged from 0 to 1. The PCM classified data were assessed for accuracy using the fuzzy error matrix (FERM) and RMSE. The accuracy statistics for RMSE values Table 4 and 5 for PCM classification of Formosat and Landsat 8 data.

Table 4. RMSE Landsat 8 data

Class	RMSE
Dense forest	0.321
Eucalyptus	0.294
Grassland	0.251
Riverine sand	0.246
Water	0.235
Wheat	0.211

Table 5. RMSE Formosat 2 data

Class	RMSE
Dense forest	0.295
Eucalyptus	0.152
Grassland	0.168
Riverine sand	0.139
Water	0.120
Wheat	0.129

The RMSE for the high resolution Formosat 2 data was lower compared to RMSE values in the moderate resolution Landsat 8 data. Land-use classes with pixels or fuzzier boundaries like dense forest and grassland, recorded high RMSE values.

5.1 Classification Results

Using artificial neural network algorithms on both Landsat 8 and Formosat 2 data, 6 file membership images corresponding to the classified classes and one hard classified image were produced. The membership files displayed the feature in a panchromatic mode (Fig 7-9) with brighter cells depicting the presence and concentration of the various land use classes.
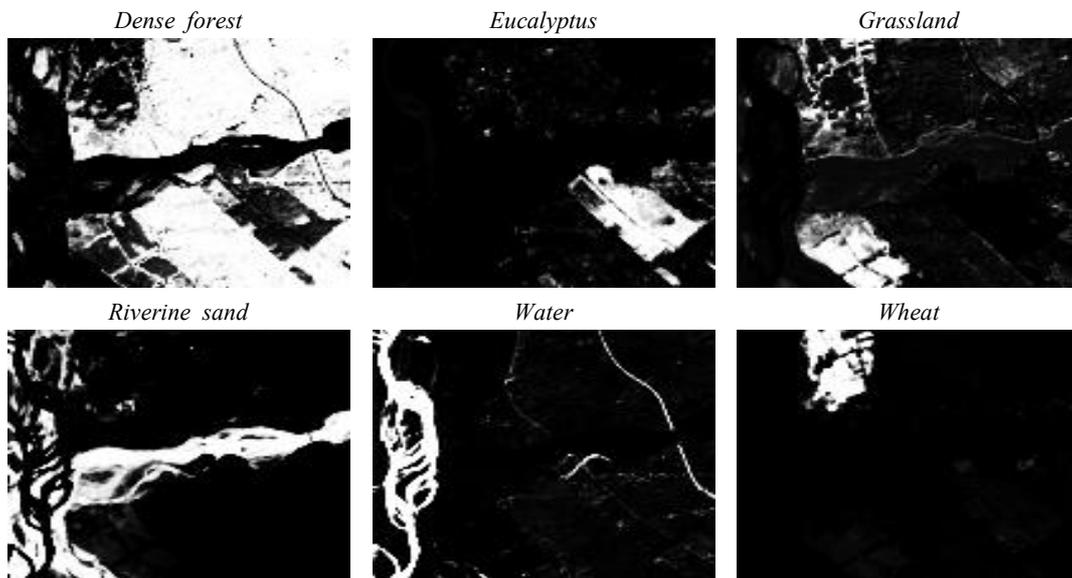
Fig. 7. Artificial Neural Network Fractional Images - Landsat 8
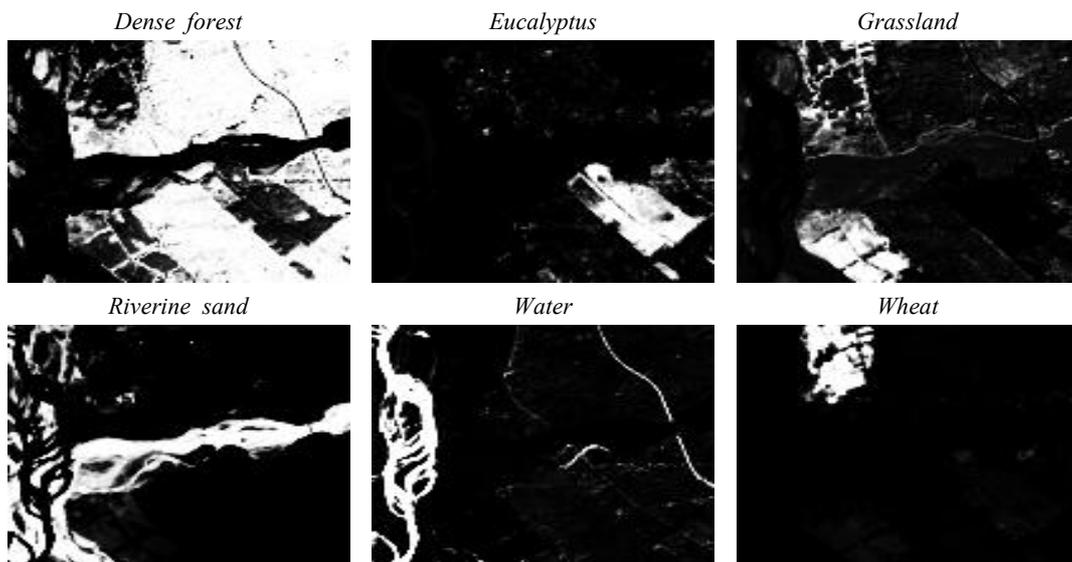
Fig. 8. Artificial Neural Network Fractional Images - Formosat 2

Dense forest	*Eucalyptus*	*Grassland*

Riverine sand	*Water*	*Wheat*
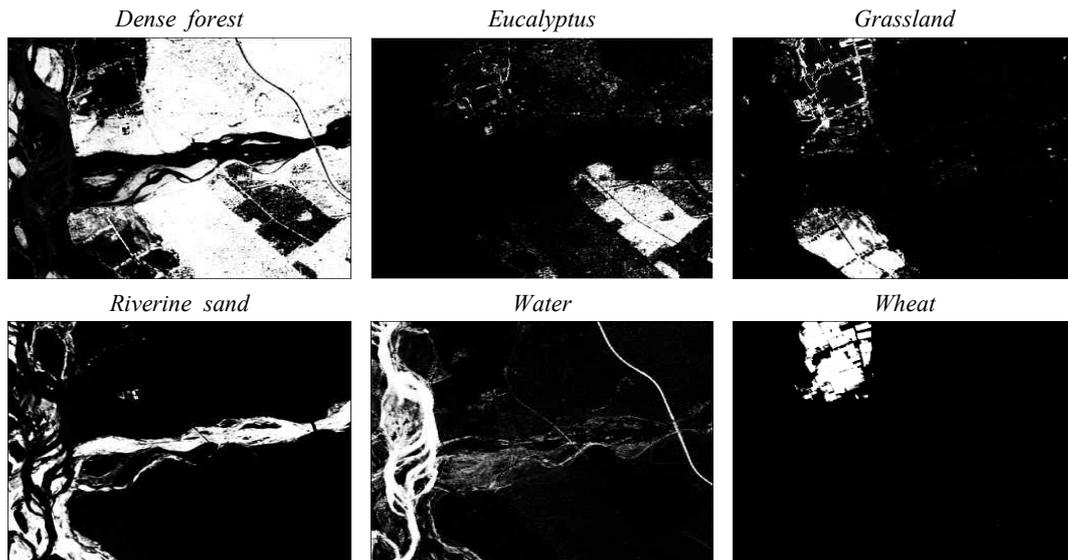
Fig. 9. Fraction images PCM-Formosat 2 (Resultant fraction images after classifying Formosat 2 with PCM)

The output from the two images did not show a greater variation in terms of feature extraction. Formosat 2 however, had a clear and distinct output because of its high spectral resolution *(8m)*.

Higher resolution imagery *(Formosat 2)* produced better results at-per-pixel level compared to moderate resolution image of Landsat 8. Subtle features like water in the Landsat 8 imagery were completely masked out by the riverine sand and could therefore not be clearly depicted compared to a similar output in Formosat 2.

Using fractional images had the advantage of de-noising the result i.e. other objects of non-interest populating the study area were masked out by the algorithm, which mainly highlighted objects/land use classes which were captured by the signature file. The panchromatic images were also represented with binary digits where 1 represented presence of the object of interest and 2 represented the masked pixels, which were represented with a dark color.

A land cover map (Fig 10) was developed to compare the output of the two algorithms from the supervised classification. Overall, the results from the two images appeared to be similar. However, Formosat 2 data produced a crisper output, especially the boundaries of farms cultivated with eucalyptus trees.
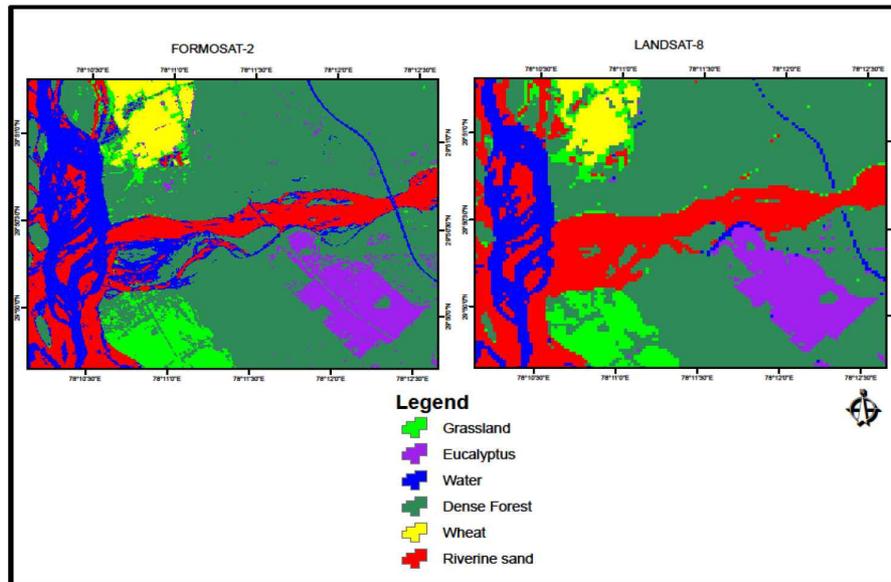
Fig. 10. Classified Formosat 2 and Landsat 8

5.2 Accuracy assessment

There is still no globally accepted method for conducting accuracy assessment for soft classified maps. Some techniques like defuzzification have been used to get hard output from soft classifiers. In any classification scheme, the more samples are collected per class, the higher the accuracy of the output. This study utilized the error matrix also called confusion matrix or contingency table, to assess the accuracy of the classification results. Error matrices compare, on a category-by-category basis, the relationship between known reference data and the corresponding results of an automated classification (Lillesand, Kiefer, & Chipman, 2004). Training set pixels classified under correct land cover categories are located along the major diagonal of the error matrix (Tables 6 & 9). All non-diagonal elements represent errors of omission or commission. Over 250 random samples were collected for this study on both Landsat 8 and Formosat 2 data.

The result from the error matrix (Table 6) shows that in all the land use classes, there were some pixels representing other classes. This was more prevalent in classes that display a similar signature pattern i.e. vegetated areas like grassland, eucalyptus and dense forest. In Landsat 8, due to the low resolution, signatures from these classes could not be separated clearly by the algorithm.

Areas covered with water also presented some similarities with riverine sand signature. This could be due to the presence of wet soil which emitted similar radiation.

Result of PCM on Formosat 2 and Landsat 8 data sets *(Table 6 to Table 8)*

Table 6. Error matrix - Landsat 8

Classified Data	Reference data						
	Grassland	Eucalyptus	Water	Dense Forest	Wheat	Riverine Sand	Row Total
Grassland	21	1	0	1	1	0	24
Eucalyptus	0	17	0	0	0	0	17
Water	0	0	15	0	0	1	16
Dense Forest	2	2	0	128	0	1	133
Wheat	0	2	0	0	8	0	10
Riverine Sand	1	0	0	0	0	55	56
Column Total	24	22	15	129	9	57	256

The accuracy totals i.e. producer vs. user's accuracy in most classes, had over 90% similarities on the Landsat data (Table 6), while Formosat 2 data (Table 9) gave accuracies of 50-60%. The greatest deviations, in terms of accuracy, on the Landsat image were observed between eucalyptus, water and wheat land use classes (Table 7). On Formosat 2 data however, grassland, water and riverine sand had a deviation of more than 30%.

The result on the two results could be attributed to various factors including the number of samples collected per class and the amount of pixels per class. Formosat 2 data, because of the high spatial resolution, could easily distinguish two land use classes within a cell. This has the impact of reducing the overall accuracy because the signatures were not collect at pixel level. The overall accuracy of the Formosat 2 data was recorded as 77.34% compared to Landsat 8 data which was 95.31%. This however, might be misleading because the summary value is basically an average and therefore doesn't reveal if error was evenly distributed between the various classes. According to (Lillesand, Kiefer, & Chipman, 2004) good results only indicate the homogeneity and spectral seperability in the training classes.

Table 7. Accuracy totals

Classified Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Grassland	24	24	21	87.50 %	87.50 %
Eucalyptus	22	17	17	77.27 %	100.00 %
Water	15	16	15	100.00 %	93 .75 %
Dense Forest	129	133	128	99.22 %	96.24 %
Wheat	9	10	8	88.89 %	80.00 %
Riverine Sand	57	56	55	96.49 %	98.21 %
Totals	256	256	244		

Result: *Overall Accuracy = 95.31%, Kappa = Observed - Expected/1 - Expected, Overall Kappa Statistics = 0.9300*

Table 8. Conditional Kappa/category

	Class Name	Feature color	Sample Points	Kappa
1.	*Grassland*	Red	108	0.8621
2.	*Eucalyptus*	Green	109	1.0000
3.	*Water*	Blue	96	0.9336
4.	*Dense Forest*	Yellow	224	0.9242
5.	*Wheat*	Cyan	72	0.7927
6.	*Riverine Sand*	Aquamarine	164	0.9770

Result of Artificial Neural Network Formosat 2 *(Table 9 to Table 11)*

The error of omission was high among the *"grassland"* and *"riverine sand"*. 14 pixels out of a total of 28 were left out and wrongly classified as *"dense forest"* (Table 9). For the riverine class. 13 were classified as *"water"* while 14 pixels were classified as *"dense forest"* making a total of 27 pixels omitted.

The highest error of commission was observed in the dense forest class. 14 pixels were classified as *"grassland"*, 6 under *"eucalyptus"*, 1 as *"water"*, andile 14 as *"riverine sand"*.

Table 9. Error matrix - Formosat 2: ANN

Classified Data	Reference data						
	Grassland	Eucalyptus	Water	Dense Forest	Wheat	Riverine Sand	Row Total
Grassland	12	0	0	1	3	0	16
Eucalyptus	0	18	0	1	0	0	19
Water	1	0	20	1	0	13	35
Dense Forest	14	6	1	106	0	14	141
Wheat	1	0	0	0	10	0	11
Riverine Sand	0	0	1	0	0	32	34
Column Total	28	24	22	109	13	59	256

Table 10. Accuracy totals: ANN

Classified Name	Reference Totals	Classified Totals	Number Correct	Producers Accuracy	Users Accuracy
Grassland	28	16	12	42.86 %	75.00 %
Eucalyptus	24	19	18	75.00 %	94.74 %
Water	22	35	20	90.91 %	57 .14 %
Dense Forest	109	141	106	97.25 %	75.18 %
Wheat	13	11	10	76.92 %	94.12 %
Riverine Sand	59	34	32	54.24 %	98.21 %
Totals	256	256	198		

Result: *Overall Classification Accuracy = 77.34%, KAPPA (K^) STATISTICS,*
Overall Kappa Statistics = 0.6796

The numbers of samples collected per land-use class were based on the spectral variability of the pixels in a given class and the geographical coverage of the class. Classes with mixed cells or noise, like areas under both water and riverine sand, so more samples for each class were collected in order to lower the error associated with pixels with similar intensity. Land-use classes like eucalyptus and wheat contained largely pure cells and therefore less samples were taken. Dense forest not only covered a sizeable area of the study area, it also had spectral variability attributed to different tree species and health status of trees under different microclimate regimes. This had the impact of misclassified classes, thus reducing the accuracy of the final product. There was a direct correlation observed between the number of samples collected per class and the Kappa coefficient value. Generally, collecting many samples for classes with pixels increases the Kappa coefficient value in Eucalyptus class (Table 12); however, classes with pure cells like wheat still yielded a high Kappa value although the numbers of samples were comparatively lower. This is because there was less noise in the wheat pixel class, which boosted the correlation between the producers and users accuracy.

Table 11. Kappa coefficient/category: ANN

№	Class Name	Feature Color	Points	Kappa
1.	*Grassland*	Green	124	0.7193
2.	*Eucalyptus*	Purple	159	0.9419
3.	*Water*	Blue	139	0.5311
4.	*Dense Forest*	Sea Green	96	0.5677
5.	*Wheat*	Yellow	91	0.9042
6.	*Riverine Sand*	Maroon	150	0.9236

Table 12. Accuracy assessment Formosat 2: Producer/User-FERM

№	Class Name	Producer Accuracy	User Accuracy
1.	*Dense forest*	92.34 %	94.66 %
2.	*Eucalyptus*	92.34 %	94.41 %
3.	*Grassland*	94.42 %	91.60 %
4.	*Riverine sand*	95.77 %	93.88 %
5.	*Water*	92.88 %	95.85 %
6.	*Wheat*	96.90 %	97.47%

The accuracy assessment results for Formosat 2 data generally, presented higher correlation between the reference data *(producer)* and classified data *(user)* in all the classes; however, higher correlation values were recorded among eucalyptus, riverine sand and wheat classes (Fig 11).
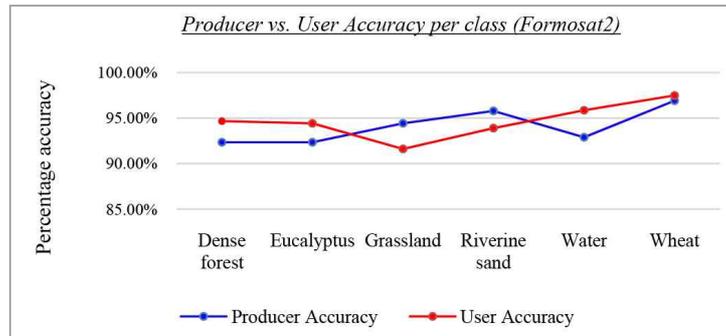
Fig. 11. Producer vs. User accuracy Formosat 2 data

Table 13. Overall categorical assessment-Formosat 2-FERM

Accuracy type	Average	Overall accuracy
Producer	94.11%	94.05%
User	94.05%	

Table 14. Accuracy assessment Landsat8 and Formosat 2-FERM

№	Class Name	Producer Accuracy	User Accuracy
1.	*Dense forest*	33.09 %	58.47 %
2.	*Eucalyptus*	31.21 %	66.07 %
3.	*Grassland*	25.98 %	16.42 %
4.	*Riverine sand*	44.27%	16.44 %
5.	*Water*	26.04 %	23.47 %
6.	*Wheat*	68.54 %	63.19 %

The correlation between producer and user accuracy was very low for Landsat 8 data; however, high values were recorded between the *"water"* and *"wheat"* classes (Table 14). Unlike other classes, these classes comprise of pure pixels when sampling is done at or near the center of the class pixels. Therefore, classified results yield values closer to the reference data. The *"eucalyptus"* class (Fig 12) yielded the lowest correlation followed by *"riverine sand"* and *"dense forest"*. The spectral signatures of these classes bore greater resemblance with the neighboring objects. The spatial resolution of a single cell also meant that two or more object classes could be mapped within a single pixel making it hard for the classification algorithm to separate each class accurately.
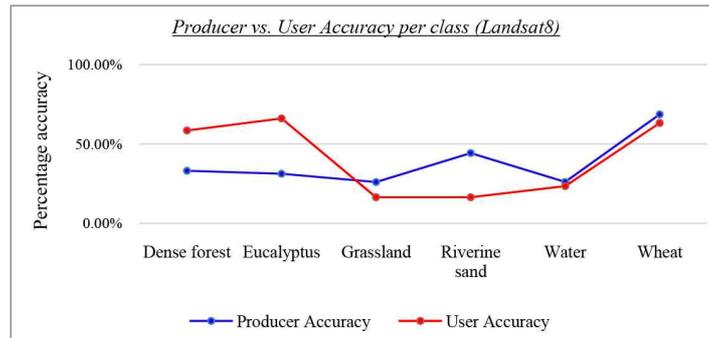
Fig. 12. Producer vs. User accuracy Landsat 8 data

Table 15. Overall categorical assessment-Landsat8 and Formosat 2-FERM

Accuracy type	Average	Overall accuracy
Producer	38.19%	50.11%
User	40.68%	

6. Summary

The two supervised classification methods used in this research study display outputs with similar results. While still in early scientific development, Artificial Neural Network shows great promise due its ability to *"recognize"* pixels and their corresponding classes once a rigorous training has been conducted.

The Possibilistic c-Means produces a more realistic and reliable result because it considers others factors like the degree of belongingness, compatibility and typicality to give a possibility of a pixel belonging to a given class. The possibilistic c-Mean therefore can handle outlier pixels far from a cluster, noise and untrained classes in an image. In the case of dense forest class, PCM was able to classify neighboring pixels with a slightly different brightness value.

In conclusion, the output observed when you combine Formosat 2 and Landsat 8 images to compute overall accuracies of the classes was lower i.e. 50.11% compared to the output from individual images i.e. 94.05% for Formosat 2 (Fig 9). The 50.11% for the overall accuracy could be attributed to the difference in resolution between the two images. This is because Landsat 8 does not give clear spectral details therefore many smaller features are generalized and their information is lost as a result of this. On the other hand, a high resolution image *(Formosat 2)* display features in detail at pixel level, while this has the advantage of presenting more information, if the algorithm is not trained at the pixel level, then there is the possibility of extracting mixed cells representing a single class. Representing a cell in one class or the other, like in hard classification would compound the errors. The strength of the PCM is that it overcomes the need to specify the number of clusters and is highly robust in the presence of noise and outliers.

References

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data (Vol. 964). US Government Printing Office.

Aqil, M., Kita, I., Yano, A., & Soichi, N. (2006). Decision support system for flood crisis management using artificial neural network. *International Journal of Intelligent Technology, 1*(1), 70-76.

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

Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction neural networks in remote sensing. *International Journal of remote sensing, 18*(4), 699-709.

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

Chattopadhyay, S., Pratihar, D. K., & De Sarkar, S. C. (2011). A comparative study of fuzzy c-means algorithm and entropy-based fuzzy clustering algorithms. *Computing and Informatics, 30*(4), 701-720.

Chawla, S. (2010). Possibilistic c-means-spatial contextual information based sub-pixel classification approach for multi-spectral data. University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede.

Debojit, B. J. H., Arora Manoj, K., & Balasubramanian, R. (2011). Study and implementation of a non-linear support vector machine classifier. International Journal of Earth Sciences and Engineering ISSN, 0974-5904.

De Jong, S. M., & Van der Meer, F. D. (Eds.). (2007). Remote sensing image analysis: including the spatial domain. Springer Science & Business Media.

Follador, M., Villa, N., Paegelow, M., Renno, F., & Bruno, R. (2008). Tropical deforestation modelling: comparative analysis of different predictive approaches. The case study of Peten, Guatemala. In *Modelling Environmental Dynamics* (pp. 77-107). Springer Berlin Heidelberg.

Foody, G. M. (1995a). Land cover classification by an artificial neural network with ancillary information. *International Journal of Geographical Information Systems, 9*(5), 527-542.

Foody, G. M. (1995b). Using prior knowledge in artificial neural network classification with a minimal training set. *Remote Sensing, 16*(2), 301-312.

Foody, G. M. (2001). Thematic mapping from remotely sensed data with neural networks: MLP, RBF and PNN based approaches. *Journal of Geographical Systems, 3*(3), 217-232.

Foody, G. M., & Cutler, M. E. (2006). Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. *Ecological modelling, 195*(1), 37-42.

Ganchimeg, G. (2015). History document image background noise and removal methods. *International Journal of Knowledge Content Development & Technology, 5*(2), 11-24.

Gong, P., Pu, R., & Chen, J. (1996). Mapping ecological land systems and classification uncertainties from digital elevation and forest-cover data using neural networks. *P. E. & R. S., 62*(11), 1249-1260.

Gong, Z., Thill, J. C., & Liu, W. (2015). ART-P-MAP neural networks modeling of land-use Change: accounting for spatial heterogeneity and uncertainty. *Geographical Analysis, 47*(4), 376-409.

Grekousis, G., Mountrakis, G., & Kavouras, M. (2016). Linking MODIS-derived forest and cropland land cover 2011 estimations to socioeconomic and environmental indicators for the European Union's 28 countries. *GIScience & Remote Sensing, 53*(1), 122-146.

Grekousis, G., Mountrakis, G., & Kavouras, M. (2015). An overview of 21 global and 43 regional land-cover mapping products. *International Journal of Remote Sensing, 36*(21), 5309-5335.

Grekousis, G., & Photis, Y. N. (2014). Analyzing high-risk emergency areas with GIS and neural networks: The case of Athens, Greece. *The Professional Geographer, 66*(1), 124-137.

Grover, N. (2014). A study of various fuzzy clustering algorithms. *International Journal of Engineering Research (IJER), 3*(3), 177-181.

Hepner, G. F., Logan, T., Ritter, N., & Bryant, N. (1990). Artificial neural network classification using a minimal training set. Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing, 56*(4), 469-473.

Hsu, K. C., & Li, S. T. (2010). Clustering spatial-temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources, 33*(2), 190-200.

Huang, W. Y., & Lippmann, R. P. (1987, June). Comparisons between neural net and conventional classifiers. In *IEEE first international conference on neural networks* (Vol. 4, pp. 485-493).

Jarvis, C. H., & Stuart, N. (1996). The sensitivity of a neural network for classifying remotely sensed imagery. *Computers & Geosciences, 22*(9), 959-967.

Kavzoglu, T. (2009). Increasing the accuracy of neural network classification using refined training data. *Environmental Modelling & Software, 24*(7), 850-858.

Krishnapuram, R., & Keller, J. M. (1996). The possibilistic c-means algorithm: insights and recommendations. *IEEE transactions on Fuzzy Systems, 4*(3), 385-393.

Krishnapuram, R., & Keller, J. M. (1993). A possibilistic approach to clustering. *IEEE transactions on fuzzy systems, 1*(2), 98-110.

Lambin, E. F., & Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences, 108*(9), 3465-3472.

Lillesand, T., Kiefer, R. W., & Chipman, J. (2014). *Remote sensing and image interpretation*. John Wiley & Sons.

Lucas, J., Freeberg, T., Krishnan, A., & Long, G. (2002). A comparative study of avian auditory brainstem responses: correlations with phylogeny and vocal complexity, and seasonal effects. *Journal of Comparative Physiology A, 188*(11-12), 981-992.

Mather, P. M. (1999). *Computer processing of remotely-sensed images: an introduction*. John Wiley & Sons.

Mather, P., & Tso, B. (2009). *Classification methods for remotely sensed data* (pp. 221-252). Boca Raton: CRC press.

Mas, J. F., & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing, 29*(3), 617-663.

Ndehedehe, C., Ekpa, A., Simeon, O., & Nse, O. (2013). Understanding the neural network technique for classification of remote sensing data sets. *NY Sci J, 6*, 26-33.

Paola, J. D., & Schowengerdt, R. A. (1995). A review and analysis of back propagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of remote sensing, 16*(16), 3033-3058.

Photis, Y. N., & Grekousis, G. (2012). Locational planning for emergency management and response: An artificial intelligence approach. *International Journal of Sustainable Development and Planning, 7*(3), 372-384.

Pratola, C., Del Frate, F., Schiavon, G., Solimini, D., & Licciardi, G. (2011, April). Characterizing land cover from X-band COSMO-SkyMed images by neural networks. In *Urban Remote Sensing Event* (JURSE), 2011 Joint (pp. 49-52). IEEE.

Suganya, R., & Shanthi, R. (2012). Fuzzy c-means algorithm-a review. *International Journal of Scientific and Research Publications, 2*(11), 1.

Stathakis D., & Vasilakos, A. (2006). Satellite image classification using granular neural networks. *International Journal of Remote Sensing, 27*(18), 3991-4003.

Thomas, B., & Nashipudimath, M. (2012). Comparative analysis of fuzzy custering algorithms in data mining. *International Journal of Advanced Research in Computer Science and Electronics Engineering, 1*(7), pp-221.

Velmurugan, T. (2012). Performance comparison between k-means and fuzzy c-means algorithms using arbitrary data points. *Wulfenia Journal, 19*(8), 234-241.

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology, 1*(1), 9-23.

[About the authors]

Dr. Ganchimeg Ganbold is a Senior Lecturer in Computer Science at the Mongolian University of Science and Technology. Her research interests include image processing, pattern recognition, remote sensing, multimedia and computer vision, traffic congestion solutions.

Mr. Stanley Chasia is an Assistant Lecturer at the Technical University of Kenya. He holds a Bachelor's degree in Geography from Moi University and a Master of Science degree in Geo-Information Systems and Remote sensing from the University of Nairobi. His research interest includes spatial modelling, hyperspectral and microwave remote sensing for environmental and earth science applications and data mining.