Journal of Research in Forestry, Wildlife & Environment Vol. 12(2) June, 2020
E-mail: jrfwe2019@gmail.com; jfewr@yahoo.com
http://www.ajol.info/index.php/jrfwe

ISBN: 2141 – 1778
Agbor et al., 2020

COMPARATIVE ANALYSIS OF NON-LINEAR ARTIFICIAL NEURAL NETWORKS AND MAXIMUM LIKELIHOOD ALGORITHMS IN FOREST COVER STUDIES

***Agbor C.F., Oluwole J. P ., Ogoliegbune,, O. M., Aigbokhan O. J. and Justina M.**

Remote Sensing & GIS Laboratory, Forestry Research Institute of Nigeria, Jericho Ibadan, Oyo State.
***Corresponding Author:** *chukwuka_friday@yahoo.com;* +234 805 544 6960

ABSTRACT

This study compares the level of uncertainty of a Back-Propagation Perceptron Network and the Maximum Likelihood classifiers (MLC) in the task of forest cover analysis. The input data comprises of bands 3, 4 and 5 of 2017 OLI Landsat image. Pixel grouping with these models was executed in Idrisi Selva using supervised technique. The degree of accuracy for each model was determined using 60 reference data. The results show that image classification with Non-linear Artificial Neural Networks algorithm (NANN), produce outputs with lower class weight RMSE of 0.02, and class weight RMSE of 0.14 was produced by Maximum Likelihood classifier. The overall accuracy of NANN (98.3%) is higher than that of MLC (80%). Standard errors at 85% confidence interval revealed NANN as a more effective statistical tool in separating forest from non-forest area. These indicate that misclassification of pixels occurred more with MLC than with NANN model. The comparison of RMSE values was possible because the same training data size, reference data and image were used for the different classifications.

Keywords: Back-Propagation, Multilayer Neural Networks, Maximum Likelihood classifier, Image classification and Misclassification.

INTRODUCTION

The assessment of landscape changes at different spatial and temporal scales requires land cover changes as one of the important parameters (Oyedotun, 2018). Several classification models have been used to assign land cover classes to pixels of remotely sensed imagery (Dengsheng *et al*, 2003, Lillesand *et al*, 2008 and Al-Ahmadi, 2008). Image classification is the process of sorting pixels into a finite number of individual classes, or categories, of data based on their pixel values (Tammy *et al*, 2013). If a pixel satisfies a certain set of criteria, then the pixel is assigned to the class that corresponds to that criterion (John *et al*, 2006). Image classification is the most used conventional land use change observation and detection method because of its ability to create series of land cover maps (El Garouani, *et al*, 2017). There are two ways to classify pixels into different categories: supervised and unsupervised classification (James, 2008, and Lillesand *et al*, 2008). Supervised classification requires the analyst to have much closer control over the classification process. In this process, you select pixels that represent patterns you recognize or can identify with help from other sources. The supervised classification methods are closely controlled by the analyst. Samples of spectral data from each feature of interest are provided for "training" the classifier to identify pixels that are spectrally similar to feature classes (Dengsheng *et al.*, 2003). Training sample data must be spectrally representative of the features of interest to effectively implement a supervised classification. The unsupervised classification is more computer-automated. Its implementation depends on the image spectral data itself to group pixels with similar spectral characteristics into the same spectral category or cluster. After classification, an analyst has the responsibility to ascertain the physical nature of each cluster and then often merges spectrally similar clusters into meaningful land-cover classes.

Good practice of image classification to determine land cover characteristics is required in environmental modeling (Al-Ahmadi *et al*, 2009). However, it is difficult to achieve these tasks in inaccessible terrain because of inadequate information of such areas. Utilizing automatic remote sensing techniques will provide a reliable solution to this problem (Al-Ahmadi *et al*, 2009). The knowledge of classification technique to use in order to achieve good result is a key to employ the right algorithm for image pixel grouping. Though there is no single "right" manner in which to approach an image classification problems (Lillesand *et al*, 2008), it is important to employ a technique that could produce a realistic feature discrimination map. This study therefore quantifies and compares the uncertainty of commonly used Maximum Likelihood Classifier and a technique that mimics the neural storage and analytical operations of the brain called Artificial Neural Networks (Non-Linear Perceptron Networks) classification technique for automatic extraction of land cover classes from Landsat OLI images.

These techniques have been applied to remotely sensed data and show good results over single data source. Related studies have been carried out in the past to compare classification algorithms based on accuracy measures usually estimated from a sample that are subject to uncertainty (Dengsheng *et al*, 2003, Al-Ahmadi *et al*, 2009, Benediktsson 2009), and one of the objectives of this study is to determine the uncertainty of the sample-based estimates,

The objective of this study is to demonstrate an approach to estimate area of a land cover class using statistics obtained for map accuracy assessment, and to construct confidence intervals that reflect the uncertainty of the area estimates obtained using Non-Linear Neural Networks and Maximum Likelihood Models.

MATERIALS AND METHODS

Study Area

The **Oluwa Forest Reserve** is located in Ondo State, Nigeria, and covers over 829 km^2 (320 sq mi).[2] It is part of the Omo-Shasha-Oluwa forest reserves, although it has become separated from the Omo and Shasha reserves (which are still connected as of 2011). The three reserves contain some of the last remaining forest in the area. Although they are biologically unique, they are threatened by logging, hunting and agriculture. Oluwa forest Reserve falls within the tropical wet-and-dry climate characterized by two rainfall peaks separated by a relatively less humid period usually in the month of August. The temperature ranged between 21 and 34°C while the annual rainfall ranged between 150 and 3000mm and mean relative humidity is 80% (Agbor *et al*, 2017).

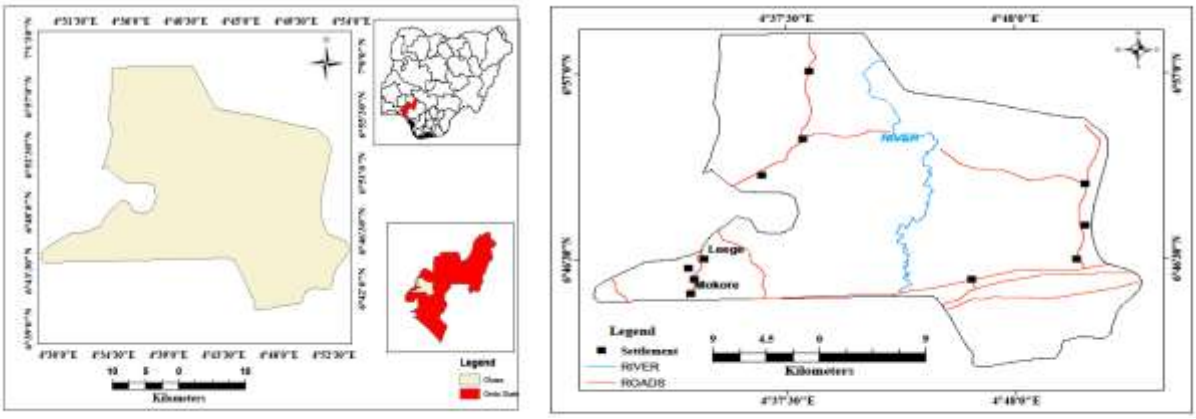

Figure 1: Study Area (Source: Ogunjemite *et al,* 2012).

Experimental Design

Image Processing Methods

This includes image preprocessing, creating spectral profile and image classification.

Image Preprocessing

Raw remote sensing data contain pixels of digital number values that correspond to a raw measure required by the sensor (Giannini *et al*, 2015). These digital numbers were converted to physical quantities to obtain quantitative information from the images. It is necessary to correct images for atmospheric effects, because the presence of the atmosphere can cause significant distortions in the radiometric signal.

Spectral Respond Pattern

Developing spectral respond pattern helps to explore what these remotely sensed images "mean" (www.clarklabs.org). To facilitate this exploration, a raster group file of the original images was first created and one of the enhanced images. This allows the link between the zoom and window actions as well as Cursor Inquiry mode across all the images belonging to the group.

There are three land-cover types that have been discerned in the image: dense forest, light forest, and non-forest. To explore how these different cover types reflect each of the electromagnetic wavelengths recorded in the original bands of imagery, reflectance values in green, red and near-infrared channels were extracted using Feature Properties query in idrisi software that allows simultaneous query of the images included in a raster image group file. Then a graph was drawn as in Figure 2. This is one of the methods used to identify features in the images (Xiao-Ling, *et al*, 2005).

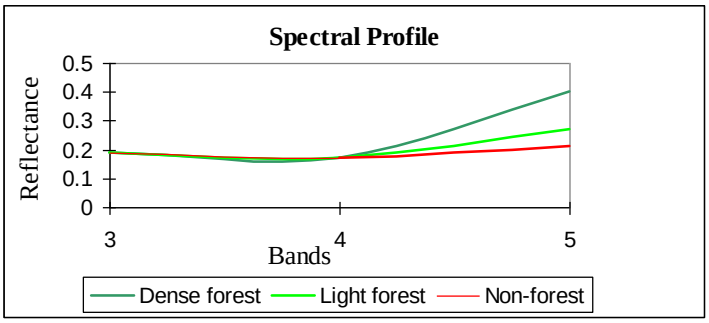

Figure 3: Spectral Respond Pattern

Image Classification with MLC and NANN

Remote sensing data classifiers of different attributes have been developed. Though it is often difficult to identify the best approach for a given study area, however, adopting a suitable classifier is of considerable importance in improving landscapes classification certainty. Different results and conclusions can be reached and this depends on the classifiers used, the study area, the image used, and training sample data available. In this study, two classifiers -maximum likelihood classifier (MLC) and Non-linear Artificial Neural Networks (NANN) algorithm were applied in the classification of Landsat image in tropical forest area, using the same training sites. The study area was classified into dense, light and non-forest areas.

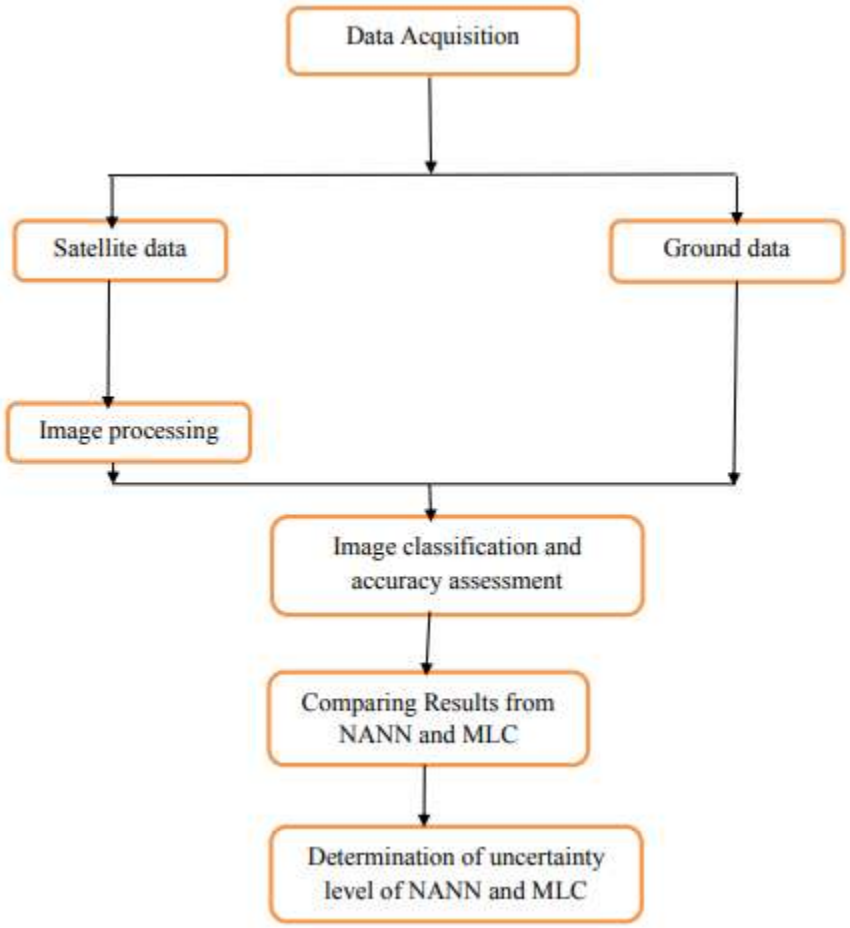

Figure 2: Flow chart of research methodology

Data Collection

To compare the effectiveness of the classifiers the study utilized Landsat images of 2017 downloaded from the official website of US Geological Survey (USGS). The study area is within the Landsat path 190 and row 55. The spatial resolution of the OLI satellite image is 30mx30m. Reference data of 60 locations were collected with Global Positioning System (GPS) for pixel training and uncertainty assessment of both classification models.

The Operation of Maximum Likelihood Classifier

Maximum Likelihood Classifier (MLC) is a supervised classification algorithm based on the Bayes theorem (Asmala *et al*, 2012). It makes use of a discriminant function to assign pixel to the class with the highest likelihood (John *et al*, 2006). Maximum Likelihood Classifier is a parametric classifier that assumes normal or near normal spectral distribution for each feature of interest. Equal prior probability among the classes is also assumed (John *et al*, 2006). To determine the class or category ω_i to which a pixel x belongs, using MLC, it is strictly the conditional probabilities $p(\omega_i/|x)$. This is the probability that the class (ω_i) is the correct category for a pixel at position x where

$$i = 1 \ldots n \qquad [1]$$

n=total number of classes.

The image classification will be performed according to equation 2.

$$x \in \omega i, \qquad [2]$$

If

$$p(\omega i \,|x) > p(\omega j \,|x) \qquad [3]$$

For all $\mathbf{j \neq I}$, where *i and j* are identified features.

This means that the pixel at position x belongs to class ω_i if $p(\omega_i \,|x)$ is larger. One major problem that associates with this classifier is that $p(\omega i \,|x)$ are

not always known. To estimate a probability distribution for a land cover type (i.e. a class) that describes the chance of finding a pixel from class $\boldsymbol{\omega i}$ at position $\boldsymbol{x}$ **i.e.** $\boldsymbol{p(x|\omega i)}$, this study ensured that sufficient training samples are available as recommended by (John *et al*, 2006). They recommend as a practical minimum that $10N$ training pixels per spectral class be used, where N is the number of channels. The dimensionality of data (images) that used in this research is low (3-channel multispectral images), therefore achieving these numbers was not be impossible.

MLH requires sufficient representative spectral training sample data for each class to accurately estimate the mean vector and covariance matrix needed by the classification algorithm. When the training samples are limited or non-representative then inaccurate estimation of the elements often results in poor classification. (Lillesand *et al*, 2008 and John *et al*, 2006).

The Operation of Non-linear Artificial Neural Networks (NANN) classifier

The performance of the Non-linear Artificial Neural Networks classifier is dependent on several factors including the quality and size of the training data sets, the complexity of the network structure and training parameters such as the learning rate (Bischof, 1998). Non-linear Artificial Neural Networks were originally designed as pattern-recognition and data analysis tools that mimic the neural storage and analytical operations of the brain (Hui *et al*, 2009 and Graciela *et al*, 2008). It utilizes multiple neural networks and is a non-linear perceptron technique that classifies remotely sensed images using backpropagation (BP) algorithm. The calculation is based on information from training sites (Dengsheng *et al*, 2003). A neural network for use in remote sensing image analysis appears as shown in figure 3, being a layered classifier composed of processing elements. (John *et al*, 2006).

It is often designed with an input layer of nodes (which has the function of distributing the inputs to the processing elements of the next layer) and an output layer from which the class labeling information is provided (Hui *et al*, 2009). In between these layers is a hidden layer which could be more in some cases. One hidden layer will be sufficient (Kanellopoulos, 1997) although the number of nodes to use in this layer is often not readily determined. The advantage of forming a classifier network is for easy handling of data sets that are not separable with a simple linear decision surface.

Training Networks Back propagation:

Before Neural Network can perform classification, the network must be trained (Graciela *et al*, 2008). This amounts to using labelled training data to help determine the weight vector w and the threshold Θ as given in equation 4 (John *et al*, 2006). In this paper efforts were being made to illustrate briefly with equations certain operations performed by this machine learning networks, for example, the calculation of weights and errors.

$$O = F(w'\varkappa + \Theta) \qquad 4$$

Multiple layer neural networks have layer structure in which successive layers of neurons are fully interconnected with connection weights controlling the strength of the connections (www.clarklabs.org). The input to each neuron in the next layer is the source of all its incoming connection weights multiplied by their connecting input neural activation value (Mustaphe *et al*, 2014).

In this study, three Landsat bands of 2017 image were used as inputs to the neural networks. These inputs were connected to the hidden layer before it came out with three classes as outputs with Θ set to zero. The output classes are dense, light, and non-forest land cover types. Neural Network architecture of 3-4-3 was used, which represents three neurons in the input layer, four neurons in the hidden layer and three in the output layer (Fig. 3).

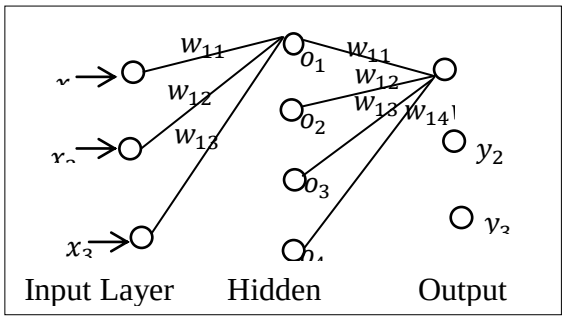

Figure 4: Structure of a 3-4-3 Network of Neurons

All the connections in figure 3 have weights in form of matrices between input and hidden layers and between hidden and the output layers. The backpropagation weight matrix from output to hidden layer and from hidden to input layer are given by equations 5 and 6 respectively (from figure 3).

$$\left.\begin{aligned} y_1 &= w_{11}o_1 + - - - - - - - - w_{14}o_4 \\ y_2 &= w_{21}o_1 + - - - - - - - - w_{24}o_4 \\ y_3 &= w_{31}o_1 + - - - - - - - - w_{34}o_4 \end{aligned}\right\} \quad [5]$$

$$\left.\begin{aligned} H_1 &= w_{11}x_1 + - - - - - - - - w_{13}x_3 \\ H_2 &= w_{21}x_1 + - - - - - - - - w_{23}x_3 \\ H_3 &= w_{31}x_1 + - - - - - - - - w_{33}o_3 \\ H_4 &= w_{41}x_1 + - - - - - - - - w_{43}o_3 \end{aligned}\right\} \quad [6]$$

Supervised training technique employed here requires known inputs (bands) and known output (classified image) (www.clarklabs.org) with output error given as in equation 7.

$$Error = target - guess\ output\ (y) \quad [7]$$

The training adjusts the weights to produce a good output, networks of 'best fit'. The error is distributed all-round the networks and this is called the back propagation. For linear perception, the "best fit" can be determined using a simple linear expression in equation 8.

$$y = M\varkappa + b \quad [8]$$

where m is the weights and b the biases. Calculating how the weights and biases change, the 'delta rule' was used (Cuiying *et al*, 2005). This determines how the weights change all round the networks from input-hidden layer to hidden-output layer (equations 9 and 10).

$$\Delta m = lr * \varkappa * error, \quad [9]$$

$$\Delta b = lr * error \quad [10]$$

In a more complex non-linear multiple neural networks as in this study the weights of the general neuron networks are given by equation 11.

$$\Delta W = \sigma(W * I + b) \quad [11]$$

Where:

$w =$ Weights, I=inputs, and b=biases.

To keep the network training stable and determine how much past outcomes should affect our weights and biases, a training parameter such as the learning rate is required (see equations 12 and 13) for weights between hidden and output layers and between input and hidden layers respectively

$$\boldsymbol{\Delta} W_{ij}^{HO} = Ir * y_{error} * (O + (1 - O)) * H^{-1} \quad [12]$$

$$\boldsymbol{\Delta} W_{ij}^{IH} = Ir * H_{error} * (O + (1 - O)) * I \quad [13]$$

In the above equations the learning rate lr is a scalar number, the error is a vector, H is the input to output layer from the hidden layer, and 0 is the output. The transformation of equation 11 into equations 12 and 13 involves computing the derivative of sigmoid of network output $s(x)$

$$S(x') = s(x) - (1 - s(x)) \quad [14]$$

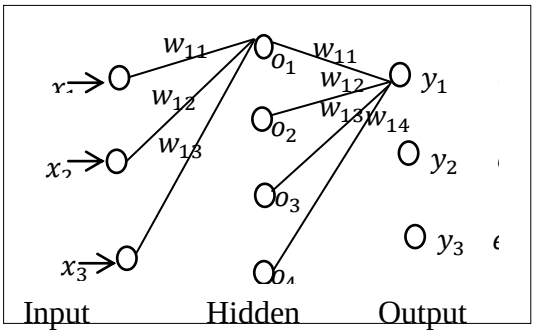

Figure 5: Error propagation Networks

Equations 15-18 show how errors are calculated based on figure 4 created for this study.

$$ErrorH_1 = \frac{w_{11}}{w_{11}+w_{12}+w_{13}w_{14}} * e_1 + \frac{w_{21}}{w_{21}+w_{22}+w_{23}w_{24}} * e_2 + \frac{w_{31}}{w_{31}+w_{32}+w_{33}w_{34}} * e_3 \quad [15]$$

$$ErrorH_2 = \frac{w_{12}}{w_{11}+w_{12}+w_{13}w_{14}} * e_1 + \frac{w_{22}}{w_{21}+w_{22}+w_{23}w_{24}} * e_2 + \frac{w_{32}}{w_{31}+w_{32}+w_{33}w_{34}} * e_3 \quad [16]$$

$$ErrorH_3 = \frac{w_{13}}{w_{11}+w_{12}+w_{13}w_{14}} * e_1 + \frac{w_{23}}{w_{21}+w_{22}+w_{23}w_{24}} * e_2 + \frac{w_{33}}{w_{31}+w_{32}+w_{33}w_{34}} * e_3 \quad [17]$$

$$ErrorH_4 = \frac{w_{14}}{w_{11}+w_{12}+w_{13}w_{14}} * e_1 + \frac{w_{24}}{w_{21}+w_{22}+w_{33}w_{24}} * e_2 + \frac{w_{34}}{w_{31}+w_{32}+w_{33}w_{34}} * e_3 \quad [18]$$

Where:
e_{1-3} Represents the error associated with each output class.
$ErrorH_{1-4}$ Represents the error from the hidden layers to each output class.

The classification algorithm based on Network of Neurons has certain advantages over parametric classification algorithms: it is non-parametric and requires little or no a priori knowledge of the distribution model of input data (Hui Yuan *et al*, 2009), the ability to estimate the non-linear relationship between the input data and desired outputs, and fast generalization capability. Studies carried out in the past on the classification of multispectral images have established that Artificial Neural Networks produces better results than traditional classification methods in terms of classification accuracy (Mustapha *et al* 2010, Benediktsson *et al* 1997, Foody *et al* 1997, Foody *et al* 1995, and Bischof *et al* 1992), though the uncertainty associated with these accuracy measures were not quantified.

Accuracy and Uncertainty Measures of MlC and Nann**s**

The strength of remote sensing is that it often allows large coverage of the area under consideration. The major weakness of this space technology is that results are never perfect, irrespective of the model used for classification. If the map has errors, then the areas of the map classes are incorrect. Therefore, the mapped areas should be adjusted for classification errors and confidence interval of area estimated. To do, the error matrix was produced for each classification. The error matrix was expressed as estimated area proportions instead of sample counts as adopted by previous studies (equation 19).

$$P_{ij=}w_i \frac{n_{ij}}{n_i} \qquad [19]$$

Where:
n_{ij} is correctly classified pixels, w_i is class weight and n_i is the sum of pixels across the columns. This gives all the information needed to estimate accuracy, area and confidence intervals. The error adjusted area of each class was computed by equation 20.

$$\acute{A}=A_{tot} \cdot r_{tot} \qquad [20]$$

A_{tot} = the area total and
r_{tot}= the row total.

Accuracy measures usually estimated from sample counts are subject to uncertainty. This uncertainty can be determined by computing the confidence interval of the estimates. A confidence interval provides a range of values for a parameter taking into account the uncertainty of the sample-based estimates. To determine the confidence interval, standard error of the area estimate was calculated as a function of area proportion and sample counts by equation 21.

$$S_e = \sqrt{\sum_{i-1}^{3} W^2 \quad \frac{\frac{n_{ij}}{n_i}\left(1-\frac{n_{ij}}{n_i}\right)}{n_i-1}} \qquad [21]$$

Where S_e is the standard error and other terms as defined earlier. This gives a standard error
of the area estimate as:

$$S(\acute{A})=A_{tot} \cdot S_e \qquad [22]$$

This in turn gives a 85% confidence intervals (Foody, 2008) as computed using equation 23.

$$\acute{A} \pm 1.44 \cdot S(\acute{A}) \qquad [23]$$

Then the accuracy measures were estimated from the area proportions (table 1)

Tabe 1: Accuracy measures of NANN and MLC

Accuracy measures	Formula	Significance/ Description
Overall accuracy:	$\sum_{j=1}^{n} P_{jj}$	The most common error estimate
User's accuracy:	$\frac{P_{ii}}{P_i}$	This is the ratio between the number of correctly classified pixels and the row total. This is necessary because users are concerned about what percentage of the classes that has been correctly classified.
Producer's accuracy:	$\frac{P_{ii}}{P_i}$	This is the ratio between the number of correctly classified pixels and the column total.

Different interpretations as to what is a good classification (table 2) have been provided by Altman D.G, (1991). The sample points used for the accuracy assessment were 60 as determined using the binomial model.

Table 2: Possible interpretation of accuracy

Accuracy range	Interpretation
< 0.2	Poor agreement
$0.2 - 0.4$	Fair agreement
$0.4 - 0.6$	Moderate agreement
$0.6 - 0.8$	Good agreement
$0.80\ to\ 1.00$	Very good agreement

Source: Altman D.G, 1991

Root Mean Square Error Estimates for the Models

This measures the difference between sample or population values predicted and the observed values. It is a measure of accuracy to compare forecasting error of different models for a particular dataset. RMSE is always non-negative and a value of 0 (zero) though almost impossible would mean a perfect fit to the data. That is, a lower RMSE is better than a higher one. Comparison of RMSE values across different datasets would be invalid because the measure depends on the scale of the numbers used. However, the comparison of RMSE values was possible here because the same training data size, reference data and the same image were used for the different classifications. The RMSE was computed using equation 23.

$$RMSE = \sqrt{\frac{1}{n}}(x_1^2 + x_2^2 + x_3^2 \dots\dots x_n^2) \qquad [23]$$

RESULTS

Land Cover distribution based on MLC and NANN**s**

The values in table 1 show the area occupied by each of the land cover types in 2017. From the table it is clear that the classifiers produced different statistics for each class. It is the main focus of this study to evaluate each classification to determine the uncertainty levels of these classification models.

Table 3: Land Cover Distribution in 2017

Land Use/Land Cover Categories	2007			
	MLC		NANN	
	Area (Km^2)	Area (%)	Area (Km^2)	Area (%)
Dense Forest	257	27	371	39
Light Forest	597	63	500	52
Non- Forest	95	10	82	9
TOTAL	953	100	953	100

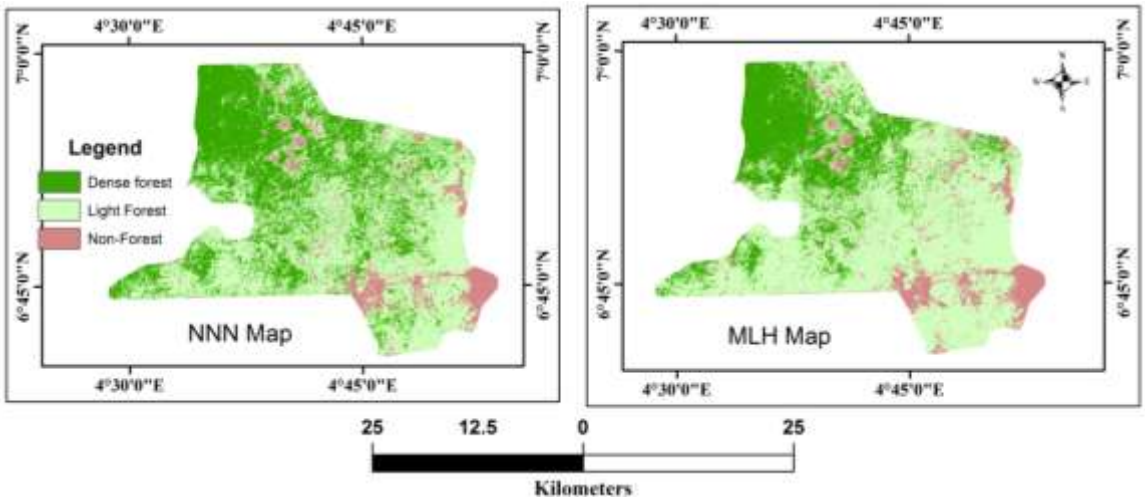

Figure 6: Output maps by both classifiers

Uncertainty Measures of MLC and NANN **Classifiers**

The sample points used for the accuracy assessment were 60 GPS point. The error matrix is presented in terms of estimated area proportions instead of absolute sample counts (table 2a and 2b). The estimated area proportions normalize the absolute sample counts by the map area and were used to calculate the users and producer's accuracy (FAO, 2016). The accuracy statistics (table 3) provides producer's accuracy (Pa), user's accuracy (Ua), overall accuracy and error-adjusted area (Pontus *et al*, 2014). The accuracy values show acceptable image classification operations using both models, with NANN model producing higher user's and producer's accuracy values. To determine error-adjusted area, the standard error for each class at 85% confidence interval was calculated from error area-estimated matrix (table 4). This reveals pixels misclassification in kilometers.

Root Mean Square Error Estimates for the Models:

The root mean square error based on the class weights and the output values (tables 5 and 6) affirm non-linear neural network classifier as a better classification algorithm than maximum likelihood classifier on remotely sensed multispectral data. Table 5 shows that image classification with Non-linear Perceptron Networks algorithm (NPN) produce outputs with class weight RMSE of 0.02, and class weight RMSE of 0.14 was produced by Maximum Likelihood classifier (table 6). RMSE of ANN is closer to zero than that of MLC.

Table 4a: Error adjusted area for MLC

MLH Classes	Dense forest	Light forest	Non-forest	Ref. data	Map area	Class weight	Error-adjusted map area
Dense forest	0.23	0.040	0	20	257	0.27	286
Light forest	0.07	0.410	0.15	26	597	0.63	429
Non-forest	0	0	0.10	14	95	0.1	238
Total	0.30	0.45	0.25	60	953	1.0	953

Table 4b: Error adjusted area for NANN

NANN classes	Dense forest	Light forest	Non-forest	Ref. data	Map area	Class weight	Error – Adjusted area
Dense forest	0.37	0.019	0	21	371	0.39	352.6
Light forest	0	0.52	0	19	500	0.52	514.6
Non-forest	0	0	0.09	20	82	0.09	85.8
Total	0.37	0.539	0.09	60	953	1.0	953

Table 5: Accuracy statistics based on error adjusted area

Class Name	MLC		NANN	
	Pa,	Ua.	Pa,	Ua.
Dense forest	85	85	100	95.2
Light forest	85	65.4	95	100
Non-forest	70	100	100	100
Overall accuracy	80%		98.3%	

Table 6: Standard error at 85% confidence interval

Landscape	S_e		$A_i \pm 1.44 \times S_e$	
	MLC	NANN	MLC	NANN
Dense forest	61	24.7	225-347	327.9-377.3
Light forest	87.7	24.7	341.2-516.6	489.9-539.3
Non-forest	51	0	187-289	82

Table 7: Class Weight Root meant Square Error of MLC

MLC classes	Class weight	Adjusted weight	Error $(x)^2$
Dense Forest	0.27	0.3	0.0009
Light Forest	0.63	0.45	0.0324
Non- Forest	0.10	0.25	0.0225
Total	20	20	0.0558

Table 6: Class Weight Root meant Square Error for NANN

ANN classes	Class weight	Adjusted weight	Error $(x)^2$
Dense Forest	0.39	0.37	0.02
Light Forest	0.52	0.54	-0.02
Non- Forest	0.09	0.09	0
Total	1.0	1.0	0.0008

DISCUSSION

Table 1 shows the classification results of NANN and MLC algorithms. Both classifiers produced more than 50% light forest and less for dense and non-forest areas. The ability of the classifiers to differentiate land cover types is closer in non-forest area than in forested areas. The accuracy statistics in table 3 provide the overall performance of the classifiers with NANN model producing higher accuracy values. The error-adjusted area table of figure 6 further shows that NANN produces better results with less error (24.7km^2 in forest areas and zero error in non-forest area) than MLC (average of 74km^2 for forest and 51km^2 for non-forest area respectively), though values from both models show acceptable image classification operations (Altman, 1991). However, image classification with Non-linear Perceptron Networks algorithm produce outputs with class weight RMSE of 0.02, and class weight RMSE of 0.14 was produced by Maximum Likelihood classifier as shown in table 6. RMSE of NANN is closer to zero than that of MLC

CONCLUSION AND RECOMMENDATIONS

In classifying remote sensing data there are several classifiers of different characteristics that have been developed. However, selecting a suitable classifier has considerable significance in improving landscapes classification certainty. Though this study shows that artificial neural network algorithm provides more accurate results than maximum likelihood classifier, both are useful in extracting information in the tropical forest area. It is important to state here that the results provide by an algorithm does not only depend on its complexity but also on the experience of the analyst. The best algorithm in the hands of an inexperience analyst will produce poor results.

REFERENCE

Agbor, C. F., Pelemo O. J., Aigbokhan, O. J., Osudiala, C. S. and Alagbe, J. (2017). Forest LossAssessment in South-West Nigeria Using Geospatial Technologies. International. *Journal of Applied Research and Technology.* 6(3): 45 – 52.

Asmala Ahmad and Shaun Quegan (2012). Analysis of Maximum Likelihood Classification on Multispectral Data. *Applied Mathematical Sciences*, 6(129): 6425 – 6436.

Bischof, H.; Schneider, W.; Pinz, A.J. (1992). Multispectral classification of landsat images using neural networks. IEEE Trans. Geosci. Remote Sens. 1992, 30, 482–490.

Bischof, H.; Leonardis, A. (1998). Finding optimal neural networks for land use classification. *IEEE Trans. Geosci. Remote Sens. 36*, 337–341.

Benediktsson, J.A.; Sveinsson, J.R. Feature extraction for multisource data classification with
artificial neural networks. Int. J. Remote Sens. 1997, 18, 727–740.

Cuiying, Z.; Liang, Z.; Xianyi, H. (2005). Classification of rocks surrounding tunnel based on improved
BP network algorithm. Earth Sci. J. China Univ. Geosci. 2005, 30, 480–486.

Canziani G., Ferrati R., Marinelli C, and Dukatz F, (2008). Artificial Neural Networks and Remote Sensing in the Analysis of the highly variable Pampean Shallow Lakes. Mathematical Biosciences and Engineering Volume 5, Number 4, October 2008. doi:10.3934/mbe.2008.5.691

Dengsheng Lu, Paul Mausel, Mateus Batistella, and Emilio Moran,, (2003). Classification Methods in the Brazilian Amazon Basin. ASPRS 2003 Annual Conference Proceedings May 2003 Anchorage, Alaska

El Garouani, A., Mulla, D. J., El Garouani, S., and Knight, J. (2017). Analysis of urban growth and sprawl from remote sensing data: Case of fez, morocco. International Journal of Sustainable Built Environment, 6, 160–169.

Food and Agriculture Organization of the United Nations Rome (2016). Map Accuracy Assessment and Area Estimation: A Practical Guide. National forest monitoring assessment working paper No.46/E

Foody, G.M. Land-cover classification by an artificial neural-network with ancillary information.
Int. J. Geogr. Inf. Syst. 1995, 9, 527–542.

Foody, G.M.; Arora, M.K. (1997). An evaluation of some factors affecting the accuracy of classification

by an artificial neural network. Int. J. Remote Sens. 1997, 18, 799–810.

F. S. Al-Ahmadi and A. S. Hames (2009). Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia. JKAU; Earth Sci., Vol. 20 No.1, pp: 167-191 (2009 A.D./1430 A.H.)James B. Campbell, (2008). Introduction to Remote Sensing, Fourth Edition

Hui Yuan, Cynthia F. Van Der Wiele and Siamak Khorram (2009). An Automated Artificial Neural Network System for Land Use/Land Cover Classification from Landsat TM Imagery. Remote Sens. 2009, 1, 243-265; doi:10.3390/rs1030243.

John A. Richards and Xiuping jia, (2006). digital image analysis. An introduction, 4th Edition.

John A. Richards · Xiuping Jia, (2006). Remote Sensing Digital Image Analysis. Springer-Verlag Berlin Heidelberg, 2006

Kanellopoulos, I.; Wilkinson, G.G. (1997). Strategies and best practice for neural network image
classification. Int. J. Remote Sens. 1997, 18, 711–725.

Lillesand, T. M. and Kiefer, R. (2008). Remote Sensing Image Interpretation. John Wiley, New York.

M. R. Mustapha, Hwee San Lim, Mohd Zubir Mat Jafri, (2010). Comparison of Neural Network and Maximum Likelihood Approaches in Image Classification. Article in Journal of Applied Sciences 10(22); 2847-2854, 2010. ISSN 1812-5654·

Oke, C. O. (2013). Terrestrial mollusc species richness and diversity in Omo Forest Reserve, Ogun State, Nigeria. *African Invertebrates* 54 (1): 93– 104.

Pontus O, Foody G.N, Stephen V, Curtis E. (2014). Good Practices for Assessing Accuracy and Estimating Area of Land Change. Remote Sensing of Environment 148 (2014) 42–57

Fitzgerald R.W. and B. G. Lees. (1992). The Application of Neural Networks to the Floristic Classification of Remote Sensing and GIS Data in Complex Terrain. Geography Department, Australian National University,

Tammy E. Parece and James B. Campbell (2013). Remote Sensing Analysis in an ArcMap Environment

T. D. T. Oyedotun (2018): Land use change and classification in Chaohu Lake catchment from multi-temporal remotely sensed images, Geology, Ecology, and Landscapes, DOI:10.1080/24749508.2018.1481657