Comparing Selected Machine Learning Algorithms in Land Use Land Cover Classification of Landsat 8 (OLI) imagery.

ABSTRACT

In recent times, there have been increase in the rate at which researchers are searching for advanced ways of carrying out land-use land-cover (LULC) mapping, especially in developing countries. This explains why supervised machine learning algorithms have become dominant methods in geo-data science. Therefore, this study is aimed at identifying different types of supervised machine learning algorithms, and their performance in LULC classification. Four machine-learning algorithms, namely Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN), and Gaussian Mixture Models (GMM) were examined. This study also attempted to validate the various models using the index-based validation method. Accuracy assessment was performed by using Kappa coefficient. The results of the LULC showed that RF classified 23% of the study area as bare land, SVM has 24% of the study area classified as bare land, K-NN also allotted 24% to bare land, while that of GMM classifier was 30%. The overall accuracy of RF, SVM, K-NN and GMM were 0.9840, 0.9780, 0.9641 and 0.9421 respectively. The Kappa Coefficient of the various classifiers were RF (0.9695), SVM (0.9580), K-NN (0.9319) and GMM (0.8916). It showed that though all the algorithms performed relatively very well, but RF performed better than the other classifiers. Finally, this study revealed that the RF algorithm is the best machine-learning LULC classifier, when compared to others. It suffices to state that, there is need for further studies since other extraneous environmental variables may be underpinning these conclusions.

Comment [DH1]: Increased rates

Comment [DH2]: Why is your spacing double? This continues further until the end.

Comment [DH3]: The readers do not need to know what supervised machine learning algorithms are dominant, we want to know about your paper and your results. Remove this.

Comment [DH4]: Unclear sentence, rewrite it clearly

Keywords: Supervised machine learning, Algorithm, Kappa Coefficient, classification

1.INTRODUCTION

Image classification defines phenomena in an image based on their spectral signatures, considered as a function wavelength. Mapping of land use land cover (LULC) dynamics has been identified as an integral part of a wide range of geospatial activities and applications [1]. Rapid and uncontrolled population growth with associated economic and industrial development, especially in developing countries with intensified LULC have become underpinning reasons for assessing changes in LULC [2,3]. Changes in LULC have a series of impacts on the environment in many ways such as increased flood, drought vulnerability, soil degradation, loss of ecosystem services, groundwater depletion, landslide hazards, soil erosion and others [4,5,6]. Over the years, researchers had deployed conventional and direct ways of mapping at various scales integrating spatial information with different levels of precision, which were laborious, time-consuming and expensive in mapping large areas [7].

Comment [DH5]: Check your usage of words. You can cut them down to be simplified.

Conversely, the satellite-based and aerial photograph-based mapping of LULC has proven cost-effective, spatially extensive, multi-temporal, and time-saving [8]. With the advancement in remote-sensing (RS) techniques and microwave sensors, satellites now provide data at various spatial and temporal scales [9,10]. Satellite images also have the advantages of multi-temporal availability as well as large spatial coverage for the LULC mapping [11,12]. In recent times, the application of machine-learning algorithms on remotely-sensed imageries for LULC mapping has been attracting considerable attention [13,14].

Therefore, researchers have been deploying various classification algorithms in the fields of Remote Sensing and Geographic Information System (GIS). They include parametric algorithms such as maximum likelihood [15], machine learning algorithms such as Random Forest RF) Artificial Neural Networks (ANNs) and Support Vector Machine (SVM) [16,17]. Machine-learning algorithms have been grouped into two categories; supervised and unsupervised techniques [18]. Examples of the supervised classification techniques include Spectral Angle Mapper (SAM), Support Vector Machine (SVM), Random Forest (RF), Mahalanobis Distance (MD), Fuzzy Adaptive Resonance Theory-Supervised Predictive Mapping (Fuzzy ARTMAP), Radial Basis Function (RBF), Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbour (K-NN), Gaussian Mixture Models (GMM), Multilayer Perception (MLP), Maximum likelihood classifier (MLC), and Fuzzy Logic [19,20].

Conversely, the unsupervised classification techniques include Affinity Propagation (AP) Cluster Algorithm, Fuzzy C-Means algorithms, K-Means algorithm, ISODATA (iterative self-organizing data) etc. [21,14]. Thus, numerous studies on the LULC modelling have been carried out using different machine-learning algorithms [22,23,24] as well as comparing the machine-learning algorithms [25,26,27,28].

It must be stated emphatically, that there are other factors apart from the type of machine learning algorithm used for LULC classification, that can affect its accuracy. Several studies found that the LULC classification using medium- resolution and low-resolution satellites do have several spectral and spatial limitations that affect its accuracy [29,30]. Though numerous studies have been conducted on land-use classification using machine-learning algorithms [1,31] but not much has been done in the comparative analysis of the various models. This study is therefore aimed at utilizing four machine-learning techniques in order to enunciate which of them can produce a high-precision LULC map based on accuracy statistics.

2. MATERIALS AND METHODS

Ileiloju/Okeigbo Local Government Area (study area) in Ondo state lies between Longitudes 6^{o} 40' and 7^{o} 14' N and Latitudes 4^{o} 38' E and 4^{o} 53' E [32]. It shares boundaries with Ondo town, Idanre and Ipetu Ijesha. In the study area, towns and villages such as Agunla, Akinsulure, Oloronba, Awopeju, Oloruntele, Bamkemo, Lisamikan and Ileoluji are notable. It covers a total area of about 698 km^2 with an average temperature of $26^{o}C$. The topography is inundated with hills such as the Ikeji and Otasun hills. The average temperature is $26^{o}C$ with a relative humidity of about 66%. The

Comment [DH6]: It is okay to have multiple paragraphs in the introduction. Break them apart.

Comment [DH7]: Some references for you:

1. https://www.sciencedirect.com/science/article/pii/S1574954121003137

2. https://www.mdpi.com/2073-445X/9/10/372

3. https://www.mdpi.com/1270896

4. https://www.tandfonline.com/doi/abs/10.1080/15481603.2021.2000350

5. https://www.sciencedirect.com/science/article/pii/S0048969722006519

6. https://www.mdpi.com/1292622

Comment [DH8]: Break this up. Hard to follow.

Comment [DH9]: This whole paragraph needs to be fixed, rewrite everything to be clear and concise. It seems that you are just writing it for the sake of filling up space.

4 topics are listed in here, can you take things slow and just choose one topic to write per paragraph? These 4 things do not even interlink with one another as part of a fluid flow.

1. Other factors leading to inaccurate classification, what are they?
2. low spatial resolution, that's very obvious, give examples of some studies which say this too?
3. There are many papers that compare machine learning algorithms, go and look harder. This is not something new, it is not novel either. In fact, many more papers are more advanced than yours. They have looked into newer forms of machine learning classifiers and AI ensembles.

study area has rivers such as Oni, Okurughu and Awo rivers flowing across the local government area in terms of the drainage system. The economy of the study area is based on the cultivation of crops such as cassava, yam, and cash crops such as oil palm, cocoa, rubber, and kola nut (https://www.manpower.com.ng>lga). It must be stated that this study did not cover the entire local government area but mainly the northern part of the local government. This was premised on the fact that the focus of this study is on the performance of different machine learning classification algorithms and not on a land use land cover change detection analysis (Figure 1).

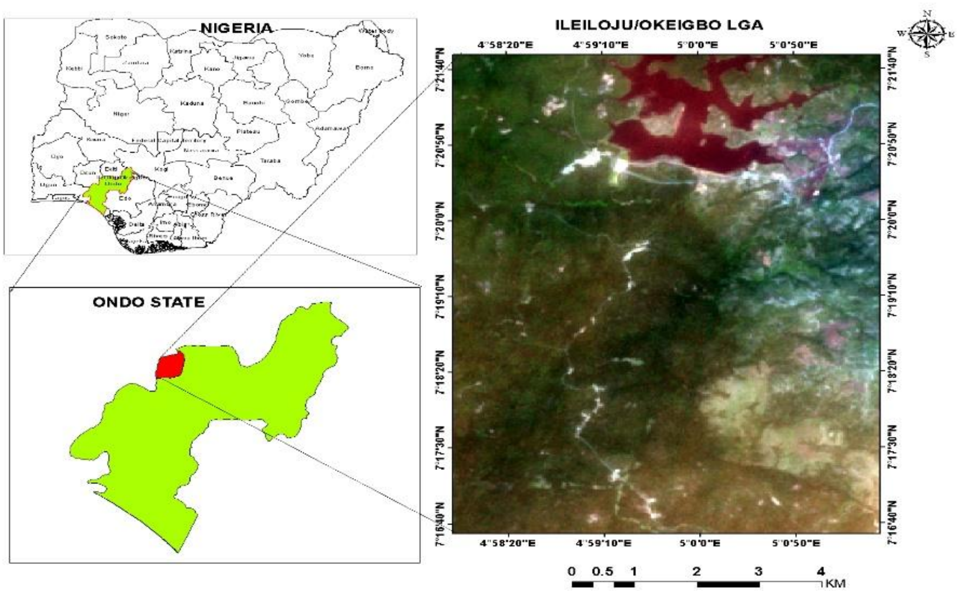

Figure 1: Map of the study area.

2.1. Materials

The Landsat 8 Operational Land Imager (OLI) image of November 25th, 2021 (path/row 190/055) was downloaded from the United States Geological Survey (USGS) website (https://earthexplorer.usgs.gov). The Google Earth image coupled with some ground control points (GCP) ~~were~~ was used ~~for the accuracy assessment of~~to assess the classified LULC maps accurately ~~the classified LULC maps~~.

2.2. Pre-processing

An atmospheric correction is a prerequisite for image pre-processing. In this study, the Dark Object Subtraction (DOS) Algorithm in QGIS 3.22 using the (SCP plugin) was deployed for the image correction. ~~Dark~~ The dark object subtraction method operates by removing the effects of scattering from the image data. It is unique ~~in the~~

Comment [DH10]: This what? This idea? This move? This decision? This choice? This map? This what?? You need to be clear.

Comment [DH11]: What band combination did you use for this image? What source of reference is the Nigeria borders from?

Comment [DH12]: Landsat 9 has been out for a while, just a comment, you could compare L8 to L9 in future works on this topic.

~~sense that~~because it derives the corrected DN (Digital Number) values majorly from the digital data without relying on outside information [33]. Dark-object subtraction (DOS) is one of the most widely used methods when it comes to reducing haze within an image. Most dark object subtraction technique assumes that there is a high probability that ~~there are at least a few pixels within an image which~~at least a few pixels within an image should be black (0% Reflectance) [33]. The (DOS) method assumes that within a satellite image, there exist features that have near-zero percent reflectance (i.e., water, dense forest, shadow), such that the signal recorded by the sensor from these features is solely a result of atmospheric scattering (path radiance), which must be removed [34,35]. ~~This study, like similar researches [36],~~Like similar research [36], this study utilized seven atmospherically corrected L8 OLI/TIRS spectral bands (Table 1).

Comment [DH13]: You need to explain what dark subtraction also does when the DN is less than 0. Dark subtraction renders some pixels unusable and their value becomes a negative or a no value. Therefore some parts of your image will have spots without pixels. What did you do to counter this?

Table 1. Landsat 8 OLI bands

Bands	Wavelength (micrometers)
Band 1 - Coastal aerosol	0.43 - 0.45
Band 2 - Blue	0.45 - 0.51
Band 3 - Green	0.53 - 0.59
Band 4 - Red	0.64 - 0.67
Band 5 - Near Infrared (NIR)	0.85 - 0.88
Band 6 - SWIR 1	1.57 - 1.65
Band 7 - SWIR 2	2.11 - 2.29

2.3. Random Forest

Random Forest (RF) is a new non-parametric ensemble machine-learning algorithm developed by Breiman [37]. It is unique ~~in the sense that~~because it can handle a variety of data, such as satellite imageries and numerical data [38]. RF is an ensemble learning algorithm premised on a decision tree, which integrates massive ensemble regression and classification trees. Several studies have shown a satisfactory performance for LULC classification using RF in the field of remote-sensing applications [13,19,27]. The higher the number of trees involved in this method the better the accuracy in the image classification and land use modelling [39,40] for instance in their study, selected 200 decision trees and submitted that the performance of RF was accurate.

2.4. Support Vector Machine

Support Vector Machine (SVM) is a non-parametric supervised machine learning method ~~aimed at solving~~to solve the binary classification problems [14]. ~~The polynomial and radial basis function (RBF) kernel in remote sensing,~~In remote sensing, the polynomial and radial basis function (RBF) kernel has been used most commonly. However,~~but~~ for LULC classification, RBF is the most popular technique, and it produces better accuracy than the other traditional methods [14]. The objective of the original SVM technique was to find the hyper-plane that can separate datasets into ~~a number of~~several classes~~, as well as to~~ and find the optimal separating hyper-plane from the available hyper-planes [41]. In this process, the vectors ensure that the width of the margin will be maximized [42]. The training samples or bordering

Comment [DH14]: Not necessarily 200 is good, sometimes 100 can do well. More is less and less is more. Put in another reference to suggest that this is not the true cause. More trees would also mean hardware problems and more human error, which leads to an offside effect in the accuracy.

samples that delineate the margin or hyper-plane of SVM are known as support vectors [20]. The operational capacity of the SVM is a function of the kernel size and density. Therefore, the differential between the simulated and the ~~real~~ actual satellite data~~, using the support vectors shows the best performance~~ shows the best performance using the support vectors [43]. The SVM was performed in QGIS 3.22 using the dzetsaka plugin.

2.4. K-Nearest Neighbour classifier

K-nearest neighbor (KNN) algorithm [37] is a method for classifying objects based on closest training examples in the feature space. K-nearest neighbor algorithm is among the simplest of all machine learning algorithms. In the classification process, the unlabeled query point is simply assigned to the label of its k-nearest neighbors. K-NN uses k-nearest neighbors from a subset of all of the training samples in determining a pixel's class or the degree of membership of a class. The selection of different values for 'K' can generate different classification results for the same sample object. KNN is a simple classification technique. KNN is used to classify the objects based on their similarity or closest training samples in the feature space [44].

2.5. Gaussian Mixture Models

A Gaussian mixture model (GMM) is useful for modeling data that comes from one of several groups. The groups might be different from each other, but data points within the same group can be well-modeled by a Gaussian distribution.

2.6. Validation of machine learning classifiers

In order to validate the results derivable from this study, the "index-based technique" has been chosen to select the best performing machine-learning technique for LULC mapping. For this purpose, three satellite-based indices; Normalized Difference Vegetation Index (NDVI), Normalized Differential Water Index (NDWI) and Normalized Difference Built-up Index (NDBI), have been classified using different threshold (equations 1-3). At the end, the area extent of the classifier-derived LULC will be statistically compared to the index-derived area extent.

$$\mathrm{NDVI} = \frac{\mathrm{NIR-Red}}{\mathrm{NIR+Red}} \quad (1)$$

$$\mathrm{NDBI} = \frac{\mathrm{MIR-NIR}}{\mathrm{MIR+NIR}} \quad (2)$$

$$\mathrm{NDWI} = \frac{\mathrm{Green-NIR}}{\mathrm{Green+NIR}} \quad (3)$$

Comment [DH15]: Arial

Comment [DH16]: Times new roman? WHY have you changed fonts

Comment [DH17]: Why is there a formula box here?

Comment [DH18]: I would suggest you use a table and put the formula and numbers inside. It is tidier that way.

2.6. Accuracy Assessment

The post-classification accuracy assessment of the LULC generated using various models has become an integral part of the classification process [45]. The Kappa coefficient statistical technique was deployed in this study ~~for assessing the level of~~to assess accuracy. Monserud and Leemans [46] suggested five levels of accuracy results: very poor (< 0.4), fair (0.4 to 0.55), good (0.55 to 0.70), very good (0.70 to 0.85) and excellent (> 0.85). Thus, the Kappa coefficient was calculated using 501 randomly selected sample points to evaluate the accuracy of LULC maps generated using different algorithms. The reference data was downloaded using ~~the~~ Google Earth Pro.

3. RESULT AND DISCUSSION

3.1. LULC Classification

In this regard, image classification based on the four advanced mathematical and machine learning algorithms including Random Forest, Support Vector Machine, K-Nearest Neighbour and the Gaussian Mixture Models. Landsat 8 (OLI/TIR) image was classified into four thematic classes: The Settlement, Bare land, Vegetation, and Waterbody. The study area is about 9,031 ha. From Table 2, out of the total area under study, RF classifier classified 392 ha (4%) as Settlement area, 2015 ha (23%) as Bare land, 6264 ha (69%) as Vegetation and 360 ha (4%) as Waterbody. The SVM classifier classified 286 ha (3%) as Settlement, 2136 ha (24%) as Bare land, 6242 ha (69%) as Vegetation and 367 ha (4%) as Waterbody. Also, 359 ha (4%) were classified as Settlement, 2153 (24%) as Bare land,6142 (68%) as Vegetation, and 378 (4%) as Waterbody by K-NN classifier. GMM classifier had 949 ha (10%) classified as Settlement, 2732 ha (30%) as Bare land,5019 ha (56%) as Vegetation and 331 ha (4%) as Waterbody. The LULC maps in Figures 2,3,4 and 5 showed that the settlement area, as classified by RF (4%), SVM (3%) and K-NN (4%) are very similar. GMM, using the same image and training samples classified 10% of the study area as settlement. With a sharp difference of about 6%, the GMM classifier tends to differ in algorithmic operations when compared to other classifiers. RF classified 23% of the study area as bare land, SVM has 24% of the study area classified as bare land. K-NN also allotted 24% to bare land, while that of GMM classifier was 30%. Vegetation thematic class has almost the same classified area extent across the four different classifiers i.e. RF (69%), SVM (69%), K-NN (68%) and GMM (56%) which is the least coverage when compared to other classifiers. Waterbody was classified as 4% by all the classifiers (Table 2).

Comment [DH19]: Back to Arial font…be consistent!

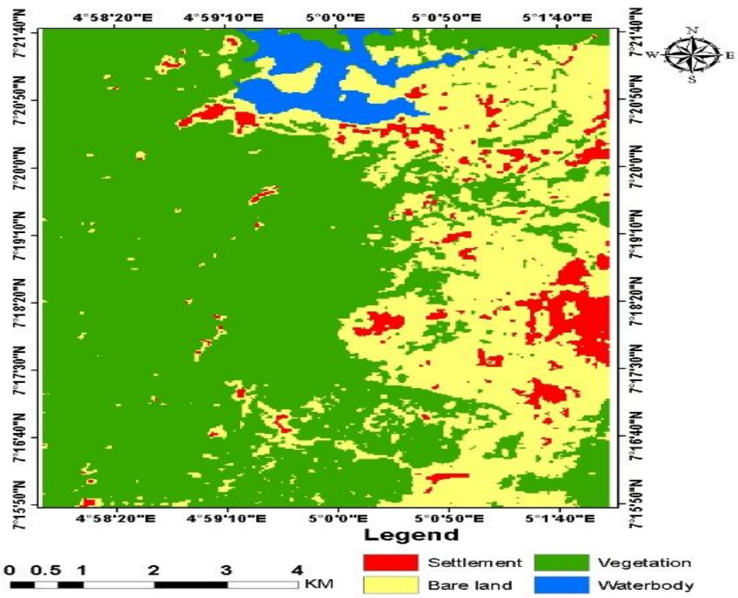

Figure 2: LULC with Random Forest (RF) classifier

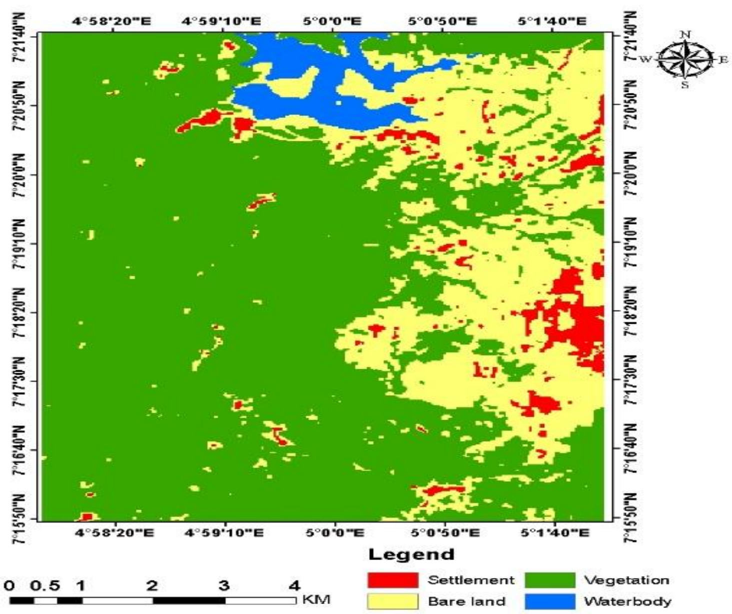

Figure 3: LULC with Support Vector Machines (SVM) classifier

Comment [DH20]: Please keep all figures at the same size. Why is SVM so small and why is KNN so flat and elongated?

Comment [DH21]: Some place markers on the map would be great. I cannot understand what I am supposed to look at. Which part of the map is a settlement and which is bare land, which corresponds to a specific named location. Put down some labels of specific towns, municipalities, etc.

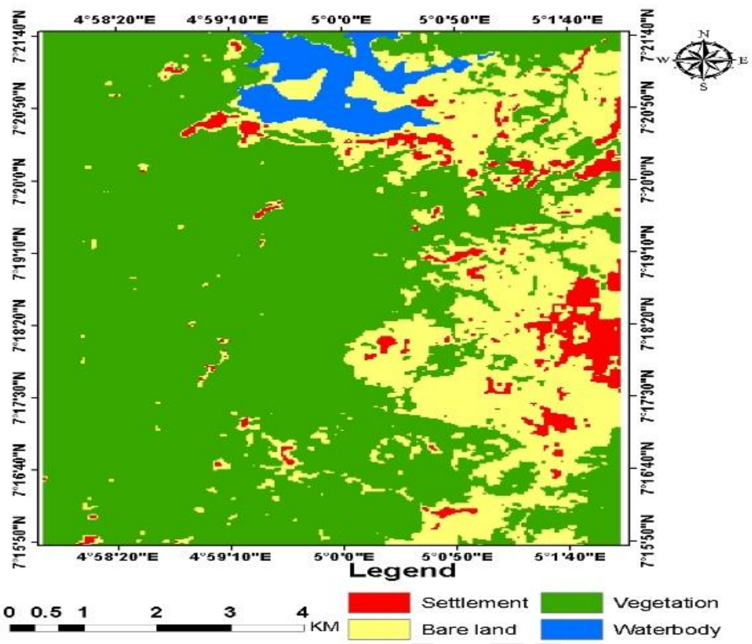

Figure 4: LULC with K-Nearest Neighbour (KNN) classifier

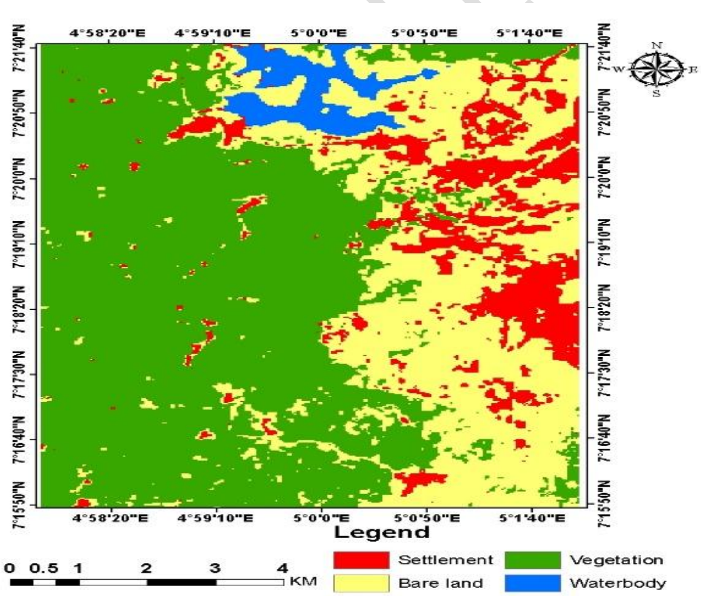

Figure 5: LULC with Gaussian Mixture Model (GMM) classifier.

Table 2 shows the percentage share of each LULC class with respect to ~~the total land coverage in the study area for each classifier~~each classifier's total land coverage in the study area.

	Random Forest (RF)		Support Vector Machine (SVM)		K-Nearest Neighbour (KNN)		Gaussian Mixture Model (GMM)	
Classes	Area(ha)	%	Area(ha)	%	Area(ha)	%	Area(ha)	%
Settlement	392	4	286	3	359	4	949	10
Bare land	2015	23	2136	24	2153	24	2732	30
Vegetation	6264	69	6242	69	6142	68	5019	56
Waterbody	360	4	367	4	378	4	331	4

Formatted Table

It is a~~n established~~ fact according to [47] that LULC classes cannot be thematically equal amongst the classification techniques, be it machine-learning algorithms or traditional classification techniques. This explained why the area extent of the various LULC classes as shown in Table 2 are different from one classifier to another. Differences in the parameter optimization of the algorithms can also be responsible the differences in area under LULC classes of different classifiers [48]. Though the studies of [13] and [27] opined that the machine-learning techniques do not have significant difference in the results, this study revealed that there could be significant differences in the LULC results of the different classifiers.

3.2. Validation of models using index-derived techniques

The results in Table 3 show the comparison between the spectral indices-derived area extent and that of the LULC derived from the classifiers. Figure 6 shows the reclassified maps of the NDVI, NDBI and the NDWI. The total area of NDBI-based is 2339 ha compared to settlement/bare land area as classified by RF classifier which is 2407 ha, with a difference of -67 ha. It shows that they are both close when compared to that of SVM (2422 ha), K-NN (2512 ha) and GMM (3681 ha) respectively. The NDVI-based vegetation area remained 6253 ha while that of RF classifier stood at 6264 ha with a difference of -11 ha. The total vegetation area extent as classified by other classifiers are SVM (6242 ha), K-NN (111 ha) and GMM (1234 ha) respectively. Water body area calculated using the NDWI was 365 ha, while that of RF classifier was 360 ha (Table 3).

Table 3. Area of LULC computed by the spectral indices and the computed areas of the LULC by the Machine Learning (ML) algorithms.

CLASS	Spectral Indices (ha)	Area (ha) computed by algorithms and their differences with spectral indices			
		RF	SVM	K-NN	GMM
Settlement/Bare land	2339	2407 (-67)	2422 (-88)	2512 (-173)	3681 (-1342)
Vegetation	6253	6264 (-11)	6242 (11)	6142 (111)	5019 (1234)

Waterbody	365	360 (5)	367 (-2)	378 (-13)	331 (34)

*Values within parenthesis indicate the difference between area computed in spectral indices and that of the algorithms.

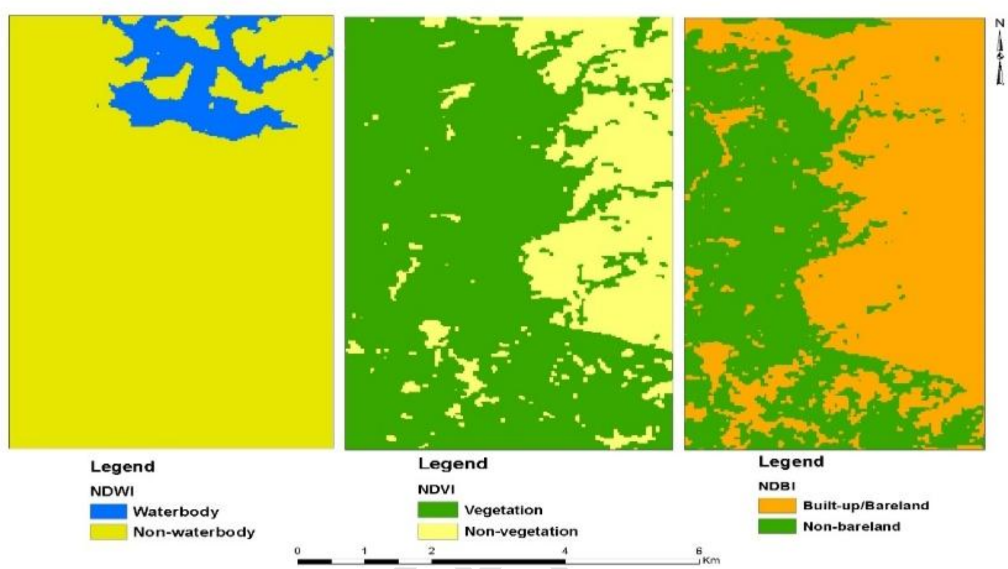

Figure 6: The index-derived maps of NDVI, NDWI and NDBI

3.3. Accuracy Assessment of the classified LULC

To validate these models' accuracy, 501 random points generated on the classified images which contain classified information. These points were then observed with the actual ground data extracted from google earth historical imagery 2021. The extracted values from the classified image vis-a-vis the reference data (google earth image) were used to calculate the error matrix, overall accuracy and Kappa coefficients of the four classifiers. Tables 4, 5, 6 and 7 showed the error matrices of the various classifiers. These error matrices helped in evaluating the performances of the various classifiers. Table 8 ~~showed~~ shows the User Accuracy and the Producer Accuracy of the various classifiers ~~in relation to~~concerning the LULC thematic classes. The producer accuracy of Settlement as classified by RF (0.9921) is the highest when compared to other classifiers, while the rest of the three classifiers (SVM, K-NN, and GMM) had approximately 0.9545. The user accuracy of the settlement class had RF (0.9167), SVM (0.9130), K-NN (0.8077) and GMM (0.6774). It showed that settlement was accurately classified by RF, but poorly classified by GMM. The results are almost the same pattern with the other classes (Table 8). The Overall Accuracy (OA) and Kappa Coefficient (K) for all the classifiers are shown in Table 9. The overall accuracy of RF, SVM, K-NN and GMM are 0.9840, 0.9780, 0.9641 and 0.9421 respectively. This was a pointer to the fact that, there was a close similarity in the performances of the classifiers in terms of OA. The Kappa

Comment [DH22]: Bigger is better

Comment [DH23]: Once again what location am I looking on this map? Map of Ileiloju/Okeigbo NDVI but where is Okeigbo on the map? Where is the town?

Comment [DH24]: Very monotonous passage. Write them clearly, all 3 sentences explain the same thing.

Comment [DH25]: Wordy, break it up and simplify.

Coefficient results of the various classifiers RF (0.9695), SVM (0.9580), K-NN (0.9319) and GMM (0.8916) showed that RF was the most accurate of all the classifiers. It suffices to state that other classifiers also performed very highly when compared to [46] Kappa Coefficient benchmark of 0.85 as excellent performance. Nevertheless, there appeared to be an excellent agreement between classified LULC map and the reality on ground. It has been found that SVM, and RF generally provide better accuracy when compared to other traditional classifiers. Some researchers have submitted that SVM and RF are the best techniques for the LULC classification compared to all other machine-learning techniques [16,19]. This study revealed that though, all the machine learning classifiers are very good in terms of LULC classification, the Random Forest is still highly recommended.

Table 4. Error matrix for RF

Error Matrix		Observed				Total
	RF	Settlement	Bare land	Vegetation	Waterbody	
Classified	Settlement	22	1	0	1	24
	Bare land	0	143	2	0	145
	Vegetation	0	0	310	1	311
	Waterbody	0	1	2	18	21
Total		22	145	314	20	501

Table 5. Error matrix for SVM

Error Matrix		Observed				Total
	SVM	Settlement	Bare land	Vegetation	Waterbody	
Classified	Settlement	21	1	0	1	23
	Bare land	0	143	2	1	146
	Vegetation	1	0	309	1	311
	Waterbody	0	1	3	17	21
Total		22	145	314	20	501

Table 6. Error matrix for K-NN

Error Matrix	Observed	Total

	K-NN	Settlement	Bare land	Vegetation	Waterbody	
Classified	Settlement	21	3	1	1	26
	Bare land	0	140	5	1	146
	Vegetation	1	0	305	1	307
	Waterbody	0	2	3	17	22
Total		22	145	314	20	501

Table 7. Error matrix for GMM

Error Matrix		Observed				Total
	GMM	Settlement	Bare land	Vegetation	Waterbody	
Classified	Settlement	21	5	4	1	31
	Bare land	0	136	8	1	145
	Vegetation	1	2	298	1	302
	Waterbody	0	2	4	17	23
Total		22	145	314	20	501

Table 8: LULC Accuracy Assessment statistics of the classifiers

	RF	SVM	K-NN	GMM

Classes	*Pa*	*Ua*	*Pa*	*Ua*	*Pa*	*Ua*	*Pa*	*Ua*
Settlement	0.9921	0.9167	0.9545	0.9130	0.9545	0.8077	0.9545	0.6774
Bare land	0.9862	0.9862	0.9862	0.9795	0.9655	0.9589	0.9379	0.9379
Vegetation	0.9872	0.9968	0.9841	0.9936	0.9713	0.9935	0.9490	0.9867
Waterbody	0.9	0.8571	0.85	0.8095	0.85	0.772727	0.8501	0.7391

Table 9. Summary of LULC Accuracy Assessment Results

Classifier	Overall Accuracy (OA)	Kappa Coefficient (K)
Random Forest	0.9840	0.9695
Support Vector Machine	0.9780	0.9580
K-Nearest Neighbour	0.9641	0.9319
Gaussian Mixture Model	0.9421	0.8916

The accuracy assessment in this study revealed an insignificant variation among the results of the classifiers. Therefore, comparing this study with some previous studies, the accuracy of LULC classification varied from one classifier to another sequel to variations in methods, techniques, time and space [49,14,27]. Variations in the classification outputs could be traceable to the influence of atmospheric, surface and illumination characteristics of the images [26]. It is pertinent to state that some other studies had reported that there are minor to moderate fluctuations in the accuracy of the LULC classification using different classifiers [50,51]. The high accuracy performance of RF classifier in this study with Kappa coefficient of 0.97 is further supported with previous studies such as that of [13] and [19] with accuracy levels 0.93 and 0.90, respectively, for the RF classifier. A small difference is found between the previous study and this study on the accuracy levels of SVM [52,53]. Furthermore, [26] noted that the accuracy of SVM and RF has very little difference, but the difference increases between either SVM and K-NN.

4. CONCLUSIONS

This study examined the accuracy of four different machine-learning classifiers for LULC classification using Landsat 8 (OLI/TIR satellite image ~~with the aim of eliciting~~to elicit the best of all the classifiers. Four different classes were identified i.e. Settlement, Bare land, Vegetation and Waterbody. The results showed that the area coverage of each LULC class varies under different classifiers. The LULC classification was subjected to an accuracy assessment analysis, using the overall accuracy and Kappa coefficient as statistical parameters for comparative analysis. At the end, the Kappa coefficient and overall coefficient showed changes in the accuracy of each LULC classifier. Both

Comment [DH26]: Discuss more on other papers in comparison to your own results. This section is lacking. You can explain in further detail what the papers have done where your research might benefit and/or if there is a similarity that you have. Potentially there are other machine learning algorithms and classifiers fro you to explore. Pixel mapping? OBIA? Hyperspectral NN Automata? Use these papers:

1. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252111

2. https://www.tandfonline.com/doi/abs/10.1080/10106049.2021.1917005

3. https://www.sciencedirect.com/science/article/pii/S1110982322000059

4. https://www.sciencedirect.com/science/article/pii/S2352938520306388

Comment [DH27]: Monotonous and repetitive.

Kappa coefficient and Overall accuracy analysis showed that RF has the highest accuracy of all classifiers applied to LULC modelling in the study area.

REFERENCES

1. Dutta D, Rahman A, Paul SK, Kundu A. Changing pattern of urban landscape and its effect on land surface temperature in and around Delhi. Environ. Monit. Assess. 2019; 191: 551. [CrossRef]

2. Hoan NT, Liou YA, Nguyen KA, Sharma RC, Tran DP, Liou CL, Cham DD. Assessing the Effects of Land-Use Types in Surface Urban Heat Islands for Developing Comfortable Living in Hanoi City. Remote Sens. 2018; 10: 1965. [CrossRef]

3. Kumari B, Tayyab M, Hang HT, Khan MF, Rahman A. Assessment of public open spaces (POS) and landscape quality based on per capita POS index in Delhi, India. SN Appl. Sci. 2019; 1: 368.

4. Cerovski-Darriau C, Roering JJ. Influence of anthropogenic land-use change on hillslope erosion in the Waipaoa River Basin, New Zealand. Earth Surf. Process. Landf. 2016; 41: 2167–2176.

5. Pal S, Kundu S, Mahato S. Groundwater potential zones for sustainable management plans in a river basin of India and Bangladesh. J. Clean. Prod. 2020; 257: 120311. [CrossRef]

6. Mahato S, Pal S. Groundwater potential mapping in a rural river basin by union (OR) and intersection (AND) of four multi-criteria decision-making models. Nat. Resour. Res. 2019; 28: 523–545.

7. Langat PK, Kumar L, Koech R, Ghosh MK. Monitoring of land use/land-cover dynamics using remote sensing: A case of Tana River Basin, Kenya. Geocarto Int. 2019.

8. Hoffmann J. The future of satellite remote sensing in hydrogeology. Hydrogeol. J. 2005; 13: 247–250. [CrossRef]

9. Scaioni M, Longoni L, Melillo V, Papini M. Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives. Remote Sens. 2014; 6: 9600–9652. [CrossRef]

10. Liou YA, Wu AM, Lin HY. FORMOSAT-2 Quick Imaging. In Optical Payloads for Space Missions; Qian SE., Ed.; Wiley: Oxford, UK, 2016; 1008, ISBN 9781118945148.

11. Wittke S, Yu X, Karjalainen M, Hyyppä J, Puttonen E. Comparison of two dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. Int. J. Appl. Earth Obs. Geoinf. 2019; 76: 167–178.

12. Viana CM, Girão I, Rocha J. Long-Term Satellite Image Time-Series for Land Use/Land Cover Change Detection Using Refined Open Source Data in a Rural Region. Remote Sens. 2019; 11: 1104. [CrossRef]

13. Adam E, Mutanga O, Odindi J, Abdel-Rahman EM. Land-use/cover classification in a heterogeneous coastal landscape using Rapid Eye imagery: Evaluating the performance of random forest and support vector machines classifiers. Int. J. Remote Sens. 2014; 35: 3440–3458.

14.Maxwell AE, Warner TA, Fang F. Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens. 2018; 39: 2784–2817.

15. Otukei JR, Blaschke T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int J Appl Earth Obs Geoinf 2010; 12: 27–31

16. Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 2011; 66(3):247–259

17. Duro DC, Franklin SE, Dubé MG. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens Environ 2012; 118: 259–272

18. Wu L, Zhu X, Lawes R, Dunkerley D, Zhang H. Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure. Int. J. Remote Sens. 2019; 40: 5973–5991.

19. Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson BA. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 2019; 152: 166–177. [CrossRef]

20. Shih HC, Stow DA, Tsai YH. Guidance on and comparison of machine learning classifiers for

Comment [DH28]: Some of the references you use are very old, like for example 8, 9, 15, 16 and so on. I would recommend that you update these references within the last 4 years. 2018 onwards. Newer references are obviously better as they can show their relevance to your study where older references might have their conclusions debunked.

Landsat-based land cover and land use mapping. Int. J. Remote Sens. 2019; 40: 1248–1274.

21. Camps-Valls G, Benediktsson JA, Bruzzone L, Chanussot J. Introduction to the issue on advances in remote sensing image processing. IEEE J. Sel. Top. Signal Process. 2011; 5: 365–369. [CrossRef]

22. Pal M. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 2005; 26: 217–222. [CrossRef]

23. Teluguntla, P.; Thenkabail, P.S.; Oliphant, A.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Huete, A. A 30-m Landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 2018, 144, 325–340.

24. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. Joint Deep Learning for land cover and land use classification. Remote Sens. Environ. 2019, 221, 173–187. [CrossRef]

25. Rogan, J.; Franklin, J.; Stow, D.; Miller, J.; Woodcock, C.; Roberts, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. Remote Sens. Environ. 2008, 112, 2272–2283. [CrossRef]

26. Li, X.; Chen, W.; Cheng, X.; Wang, L. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. Remote Sens. 2016, 8, 514. [CrossRef]

27. Camargo, F.F.; Sano, E.E.; Almeida, C.M.; Mura, J.C.; Almeida, T. A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images. Remote Sens. 2019, 11, 1600. [CrossRef]

28. Jamali, A. Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: A case study of the northern region of Iran. SN Appl. Sci. 2019, 1, 1448. [CrossRef]

29. Manandhar, R.; Odeh, I.O.; Ancev, T. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sens. 2009, 1, 330–344. [CrossRef]

30. Yang, C.; Wu, G.; Ding, K.; Shi, T.; Li, Q.; Wang, J. Improving land use/land cover classification by integrating pixel unmixing and decision tree methods. Remote Sens. 2017, 9, 1222. [CrossRef]

31. Pal, S.; Ziaul, S.K. Detection of land use and land cover change and land surface temperature in English Bazar urban centre. Egypt. J. Remote Sens. Space Sci. 2017, 20, 125–145. [CrossRef]

32. Adeyeye O.R, Adefusisoye A.A, Awokunmi E.E and Olanipekun E.O. (2016). Evaluation of polycyclic aromatic hydrocarbons in water from hand dug wells at Ile-Oluji, Nigeria. *IOSR Journal of Environmental Science, Toxicology and Food Technology*. 10 (9), 112-119

33. Chavez Jr, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction for multispectral data. Remote Sensing of Environment 24, 459-479

34. Paul, M. Mather; and Brandt, Tso; "Classification Methods for Remotely Sensed Data", Second Edt., CRC Press, Taylor and Francis Group, LLC, pp(5-7, 53-54), (2009).

35. Michael, Köhl; Steen, S. Magnussen; and Marco, Marchetti; "Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory", Springer, New York, pp (205-207), (2011).

36. Karakizi C, Vakalopoulou M, Karantzalos K (2017) Annual croptype classifcation from multitemporal landsat-8 and sentinel-2 data based on deep-learning. In: Proceedings of the 37th international symposium on remote sensing of environment (ISRSE37), 2017, Tshwane, South Africa

37. Breiman L (2001) Random forests. Mach Learn 45(1):5–32

38. Abdullah, A.Y.M.; Masrur, A.; Adnan, M.S.G.; Baky, M.; Al, A.; Hassan, Q.K.; Dewan, A. Spatio-temporal patterns of land use/land cover change in the heterogeneous coastal region of Bangladesh between 1990 and 2017. Remote Sens. 2019, 11, 790.

39. Liaw, A.; Wiener, M. Classification and Regression by randomForest. R News 2002, 2, 18–22. 85.

40. Feng, Q.; Gong, J.; Liu, J.; Li, Y. Flood mapping based on multiple endmember spectral mixture analysis and random forest classifier—The case of Yuyao, China. Remote Sens. 2015, 7, 12539–12562. [CrossRef]

41. Srivastava, P.K.; Han, D.; Rico-Ramirez, M.A.; Bray, M.; Islam, T. Selection of classification techniques for land use/land cover change investigation. Adv. Space Res. 2012, 50, 1250–1265. [CrossRef]

42. Bouaziz, M.; Eisold, S.; Guermazi, E. Semiautomatic approach for land cover classification: A remote sensing study for arid climate in Southeastern Tunisia. Euro Mediterr. J. Environ. Integr. 2017, 2, 24. [CrossRef]

43. Mathur, A.; Foody, G.M. Multiclass and binary SVM classification: Implications for training and classification users. IEEE Geosci. Remote Sens. Lett. 2008, 5, 241–245. [CrossRef]

44. Olmanson.L.G.; Bauer.M.E.; Brezonik.P.L. A 20 Year Landsat Water Clarity Census of Minnesota's 10,000 Lakes. Remote Sens. Environ. 2008, 112, 4086–4097.

45. Hurskainen, P.; Adhikari, H.; Siljander, M.; Pellikka, P.K.E.; Hemp, A. Auxiliary datasets improve accuracy of object-based land use/land cover classification in heterogeneous savanna landscapes. Remote Sens. Environ. 2019, 233, 111354. [CrossRef]

46. Monserud, R.A.; Leemans, R. Comparing global vegetation maps with the Kappa statistic. Ecol. Model. 1992, 62, 275–293. [CrossRef]

47. Abdi, A.M. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GISci. Remote Sens. 2019, 1–20. [CrossRef]

48. Erbek, F.S.; Özkan, C.; Taberner, M. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. Int. J. Remote Sens. 2004, 25, 1733–1748. [CrossRef]

49. Rodriguez-Galiano, V.F.; Chica-Rivas, M. Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. Int. J. Digit. Earth 2014, 7, 492–509. [CrossRef]

50. Islam, K.; Jashimuddin, M.; Nath, B.; Nath, T.K. Land use classification and change detection by using multi-temporal remotely sensed imagery: The case of Chunati wildlife sanctuary, Bangladesh. Egypt. J. Remote Sens. Space Sci. 2018, 21, 37–47. [CrossRef]

51. Leyk, S.; Uhl, J.H.; Balk, D.; Jones, B. Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States. Remote Sens. Environ. 2018, 204, 898–917. [CrossRef] [PubMed]

52. Szuster, B.W.; Chen, Q.; Borger, M. A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones. Appl. Geogr. 2011, 31, 525–532. [CrossRef]

53. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing machine learning classifiers for object-based land cover classification using very high-resolution imagery. Remote Sens. 2015, 7, 153–168. [CrossRef]