



## Assessing Image Classification Accuracy with Principal Component Analysis Algorithm Case Study: Odeda LGA of Ogun State, Southwest Nigeria

AIGBOKHAN, OJ\*; ESSIEN, NE; OGOLIEGBUNE, OM; AFOLABI, OS; ADAMU, IS

*Environmental Modeling and Biometrics Department, Forestry Research Institute of Nigeria, P.M.B 5054 Jericho, Ibadan, Oyo State, Nigeria*

\*Corresponding Author Email: [oseyomon255@gmail.com](mailto:oseyomon255@gmail.com); Tel: +2348054039119

Other Authors Email: [essiensi2018@gmail.com](mailto:essiensi2018@gmail.com), [mogoliegbune@gmail.com](mailto:mogoliegbune@gmail.com), [mosadioluwaolamidipupo@mail.com](mailto:mosadioluwaolamidipupo@mail.com), [lollysbergally15@gmail.com](mailto:lollysbergally15@gmail.com)

**ABSTRACT:** The aim of this study is to assess image classification accuracy using the instrumentality of Principal Component Analysis (PCA). It is focused on evaluating the accruable benefits of Principal Component Analysis as part of an image preprocessing procedure for image classification. Land-use-land-cover (LULC) and accuracy assessment datasets were obtained with remote sensing and geographic information system's software. The principal component analysis was statistically used to assess the level of correlation amongst bands in Landsat 8. The image classification was premised on the Maximum Likelihood classifier for land use land cover analysis. To ascertain the accuracy of the classified images, the Producer's accuracy, User's accuracy and Kappa coefficient derivatives of accuracy assessment was calculated. The results revealed that the first three PCs of the raw Landsat data accounted for 99.37 % variance of the original Landsat data, while the last three PCs represented only 0.63% of the original data. The results of land use land cover based on raw bands composite were Forest (41%), Shrubs (33%) and Built-up (26%) respectively. On the other hand, land use land cover based on Principal Component Analysis showed Forest (39%), Shrubs (39%) and Built-up (22%) respectively. Comparing the results of Kappa coefficients of both LULC of raw bands' composite was 0.88 while that of PCA was 0.91. Conclusively, there is a significant level of difference in the classification outputs of PCA derived classification and that of raw Landsat bands' composite.

DOI: <https://dx.doi.org/10.4314/jasem.v26i5.11>

**Open Access Article:** (<https://pkp.sfu.ca/ojs/>) This an open access article distributed under the Creative Commons Attribution License (CCL), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Impact factor:** <http://sjifactor.com/passport.php?id=21082>

**Google Analytics:** <https://www.ajol.info/stats/bdf07303d34706088ffffbc8a92c9c1491b12470>

**Copyright:** © 2022 Aigbokhan *et al*

**Dates:** Received: 08 December 2021; Revised: 13 April 2022; Accepted: 11 May 2022

**Keywords:** Maximum Likelihood, Kappa coefficient, LULC, Principal Component Analysis

The dimensionality of a data set can be mathematically reduced using the instrumentality of Principal Component Analysis (PCA) (Munyati 2004). The reduction in the dimensionality of a data set can further accentuate the visual characteristics of digital images. In sequel, it makes data processing and analysis more concise and manageable. There is fusion of PCA method with the digital image processing of satellite images with the capability of reducing a number of correlated image information bands to few uncorrelated bands (David 2017, Estornell *et al.* 2013). Researchers from different fields of human endeavor have successfully used PCA model in their studies (Gasmi *et al.* 2016; Lan *et al.* 2017; Zhao *et al.* 2017; Marchetti *et al.* 2020). For instance, Li *et al.* 2020 proposed the use PCA and high-dimensional model representation to estimate the probabilistic

power flow; Geng *et al.* (2020) used the PCA based memory network to predict short-term wind speed; Schwartz *et al.* (2020) proposed a PCA method to conduct change detection in radar images; Gao, *et al.* (2019) worked on extreme learning machine and adaptive PCA methods for network intrusion. Remote sensing is the art or science of obtaining and analyzing information about phenomenon, area or object using a physical device without a physical contact (Jeeva and Naraana 2016). Object classification can be performed through the spectral analysis of the reflected or emitted radiant energy of the target (Meera *et al.* 2015). Oftentimes, remote sensing deals with multispectral images with highly correlated bands. In other to save data storage space and computing time, such bands could be combined into new, less correlated images by PCA. Multi-dimensional Principal Component works

\*Corresponding Author Email: [oseyomon255@gmail.com](mailto:oseyomon255@gmail.com)

directly on the vector data of digital image where each band is taken as dimension of the matrix. The work is done on the principle of applying PCA and the methodology had been tested on several standard images (Dwivedi *et al.* 2006). Many researchers are exploring new scientific ways of improving on the performance of PCA in image analysis. To improve the performance of image compression, extended PCA based method to can be utilized to compress single image rather than a set of separated images. This method uses the correlations between three color components of an image (Mofarreh *et al.* 2015). PCA approach for identification and analysis of multi-layer images present comparatively better results than previously used techniques (Imran *et al.* 2005). The most common feature-extraction method is PCA, which transforms the data into a new set of principle components (PCs) that describes the underlying structure of the original dataset (Zhang and Mishra 2012). Multi-collinearity is simply a high degree of correlation among predictive variables in multiple regressions (Klainbaum *et al.* 1998). One of the ways of solving the problem of multi-collinearity is the application of Principal Component Analysis (PCA). Principal Component Analysis is a traditional multivariate statistical method commonly used to reduce the number of predictive variables and solve the multi-collinearity problem (Bair *et al.* 2006). Accuracy assessment of Land Cover maps, produced from remotely sensed data, involves comparing thematic maps with reference data (Congalton 1991). Since there were no suitable existing reference data that could be used for all locations on the earth's surface, a practical and statistically sound sampling plan was designed by Zhu *et al.* (2000) to characterize the accuracy of common and rare classes for the map product using National Aerial Photography Program (NAPP) photographs as the reference data. The sampling design was developed based on the following criteria: (1) ensure the objectivity of sample selection and validity of statistical inferences drawn from the sample data, (2) distribute sample sites spatially across the region to ensure adequate coverage of the entire region, (3) reduce the variance for estimated accuracy parameters, (4) provide a low-cost approach in terms of budget and time, and (5) be easy to implement and analyze (Zhu *et al.* 2000). The need for assessing the accuracy of a map generated from any remotely sensed data has become universally recognized as an integral project component. In the last few years, most projects have required that a certain level of accuracy be achieved for the project and map to be deemed a success (Ross and John, 2004). Therefore, the objective of this paper is to The aim of this study is to assess image classification accuracy using the instrumentality of Principal

Component Analysis in Odeda LGA of Ogun State, Southwest Nigeria

## MATERIALS AND METHODS

The study area, Odeda Local Government is one of the twenty (20) Local Governments Areas in Ogun State, south west Nigeria. It is located between 7°13' and 7°30' N of latitude, 3°11' and 3°46' E of longitude (Fig. 1) and covers a total land area of about 1,560 km<sup>2</sup>. It has a population of 109,449 according to the 2006 population census (NPC, 2006). The study area is predominantly rural with about 25-30 semi-urban areas and 860 villages and hamlets (Adedeji *et al.* 2020).

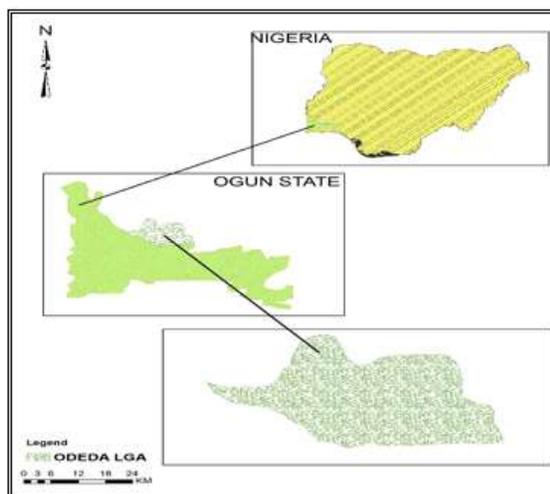


Fig. 1 Map of Odeda Local Government Area in Ogun State, Nigeria

The Landsat imagery used for this study was downloaded from the official website of Global Land Cover Facility (GLCF) – (<http://www:glcf.umiacs.umd.edu>). Satellite imagery of March 16<sup>th</sup>, 2021 from Path 191 and Row 055 was used. Pre-processing of image helps to enhance and improve the quality of the image (Mussie 2011). Radiometric and geometric corrections were performed on the image to enhance output quality. Image preprocessing was carried out using nearest neighbor interpolation algorithm. When compared to other interpolation algorithms such as Linear interpolation, Bilinear interpolation and Bi-cubic interpolation techniques, Nearest Neighbor interpolation is quite simple and faster to calculate. Nearest neighbor interpolation method assigns each interpolated output pixel value of the nearest sample point in the input image. The interpolation kernel for the nearest neighbor is represented in equations 1 and 2 (Venkata 2019).

$$h(x) = \begin{cases} 0 & / x / > 0 \\ 1 & / x / < 0 \end{cases} \quad (1)$$

Where  $x$  is the pixel value?

The vector map of the study area was used to clip out Landsat 8 (OLI) image. Bands 2, 3, 4, 5, 6 and 7 which were stacked for further processing in ArcGIS 10.4. Table 1 shows the band statistics of the various bands that were used for this study. The clipped Landsat bands 2, 3, 4, 5, 6, and 7 were subjected to Principal Component Analysis as contained in the Spatial Analyst Tool, of ArcGIS software. The derived PC<sub>1</sub>, PC<sub>2</sub> and PC<sub>3</sub> were composited as a dataset for image classification. Another Landsat raw bands dataset were also composited for another classification, which was used for comparative analysis.

**Table 1** Landsat 8 (OLI) Bands Statistics

Layers	Minimum	Maximum	Mean	Std. Deviation
Band 2	9705	12592	10152.52	200.29
Band 3	8777	12790	9378.46	263.31
Band 4	7958	13869	8907.84	520.99
Band 5	7837	21383	14103.34	827.73
Band 6	6536	25510	12910.80	1270.91
Band 7	5956	54562	9610.32	1281.32

Principal Component analysis (PCA) is a statistical method used to decrease a set of correlated multivariate measure to a narrower set where the characteristics are uncorrelated. (Venkata *et al.* 2019). It is one of the most common feature-selection methods, which transforms the data into a new set of principle components (PCs) that describes the underlying structure of the original dataset (Zhang and Mishra 2012). PCA method have been integrated in the digital image processing of satellite images as a unique conversion in which a number of correlated image information bands have been reduced to few uncorrelated bands (David 2017). Mathematical and statistical concepts used to calculate PCA are: standard deviation, covariance, eigenvalues, eigenvectors and linear transformations as shown in equation 2 (Masoumeh *et al.* 2016).

$$X_{n,b} = \left[ \begin{pmatrix} x_{2,1} & \dots & x_{2,n} \\ \vdots & \ddots & \vdots \\ x_{7,1} & \dots & x_{7,n} \end{pmatrix} \right] \quad (2)$$

Where;  $n$  represents number of pixels and  $b$  stands for the number of bands.

The matrix (equation 3) can be simplified, considering each group as a vector.

$$X_k = \begin{bmatrix} x_2 \\ x_3 \\ \cdot \\ \cdot \\ x_7 \end{bmatrix} \quad (3)$$

Where;  $k$  is the band number.

The covariance matrix's Eigenvalues must be calculated to decrease the dimensionality of the original bands and covariance matrix can be calculated as shown in equation 4

$$\sigma_{i,j} = \frac{1}{N-1} \sum_{p=1}^N (DN_{p,i} - \mu_i)(DN_{p,j} - \mu_j) \quad (4)$$

Where;  $\sigma_{i,j}$  is the covariance of the different bands of each pair;  $DN_{p,i}$  is a digital number of a pixel  $p$  in the band  $i$ ,  $DN_{p,j}$  is a digital number of a pixel  $P$  in the band  $j$ ,  $\mu_i$  and  $\mu_j$  are the averages of the DN for the bands  $i$  and  $j$ , respectively.

From the variance-covariance matrix (equation 2), the eigenvalue ( $\lambda$ ) are calculated as the roots of the characteristic (equation 5) (Masoumeh *et al.*, 2016).

$$\det (C-\lambda I) = 0 \quad (5)$$

Where,  $C$  is the covariance matrix of the bands and  $I$  is the diagonal identity matrix.

The matrix of the Principle Component can be described as shown equation 6

$$y = \begin{bmatrix} w_{2,1} & w_{2,n} \\ \cdot & \cdot \\ \cdot & \cdot \\ w_{7,1} & w_{7,n} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ \cdot \\ \cdot \\ x_7 \end{bmatrix} \quad (6)$$

Where;  $y$  is the principal component vector,  $w$  is the transformation matrix and  $x$  is the original data vector

In this study, the Maximum Likelihood (ML) supervised classification method was used. It is derived from the Bayes theorem, which states that a posteriori distribution  $P(i|\omega)$ , i.e., the probability that a pixel with feature vector  $\omega$  belongs to class  $i$ , (Asmala and Shaun 2012) as shown in equation 7.

$$P(i|\omega) = \frac{P(i|\omega)P(\omega)}{P(\omega)} \quad (7)$$

Where;  $P(\omega|i)$  is the likelihood function,  $P(i)$  is the a priori information, i.e., the probability that class  $i$  occurs in the study area and  $P(\omega)$  is the probability that  $\omega$  is observed, which can be written as:

$$P(\omega) = \sum_{i=1}^m P(\omega | I_i) P(I_i) \tag{8}$$

Where M is the number of classes; P(ω) is often treated as a normalization constant to ensure  $\sum_{i=1}^m P(I_i | \omega)$  sums to 1 (Asmala and Shaun 2012)

In ML classification, each class is enclosed in a region in multispectral space where its discriminant function is larger than that of all other classes. These class regions are separated by decision boundaries, where, the decision boundary between class i and j occurs when:

$$g_i(\omega) = g_j(\omega) \tag{9}$$

Accuracy assessment is a general term for comparing the classification of geographical data that are assumed to be true (reference), in order to determine the accuracy of the classification process. Error matrix has become a standard in the accuracy assessment of remote-sensing classification results (Nagamani 2015). The most common means of reporting the reliability of a land cover map derived from satellite data is the error or confusion matrix (Table 2), also called a contingency table (Congalton 1991). The error or confusion matrix represents a tabulated error made in a classification. The columns stand for categories on the ground while the rows represent the categories assigned in the mapping project. The overall accuracy represents the sum of the diagonal elements divided by the total number of pixels in the table. The producer's accuracy is calculated by dividing the number of pixels accurately classified in a given category by the total number of pixels of that category that were sampled on the ground. When the number of pixels in a category that were correctly classified is divided by the total number of pixels that were assigned to that category in the classification, the user's accuracy is obtained. The Kappa coefficient and the error matrix are considered as common techniques in measuring

the accuracy of thematic maps generated by the classification process. The Kappa coefficient can be calculated using equation 10

$$K_{APPA} = \frac{N \sum^k X_{ii} \sum^k (X_{i+} \times X_{+i})}{N^2 - \sum^k (X_{i+} \times X_{+i})} \tag{10}$$

Where; KAPPA = Kappa index, k = number of matrix files, Xii = observation number on row i and column I (along the diagonal), (X<sub>i+</sub> and X<sub>+i</sub>) = total marginal for row i and column i, respectively, N = total number of observations.

## RESULTS AND DISCUSSION

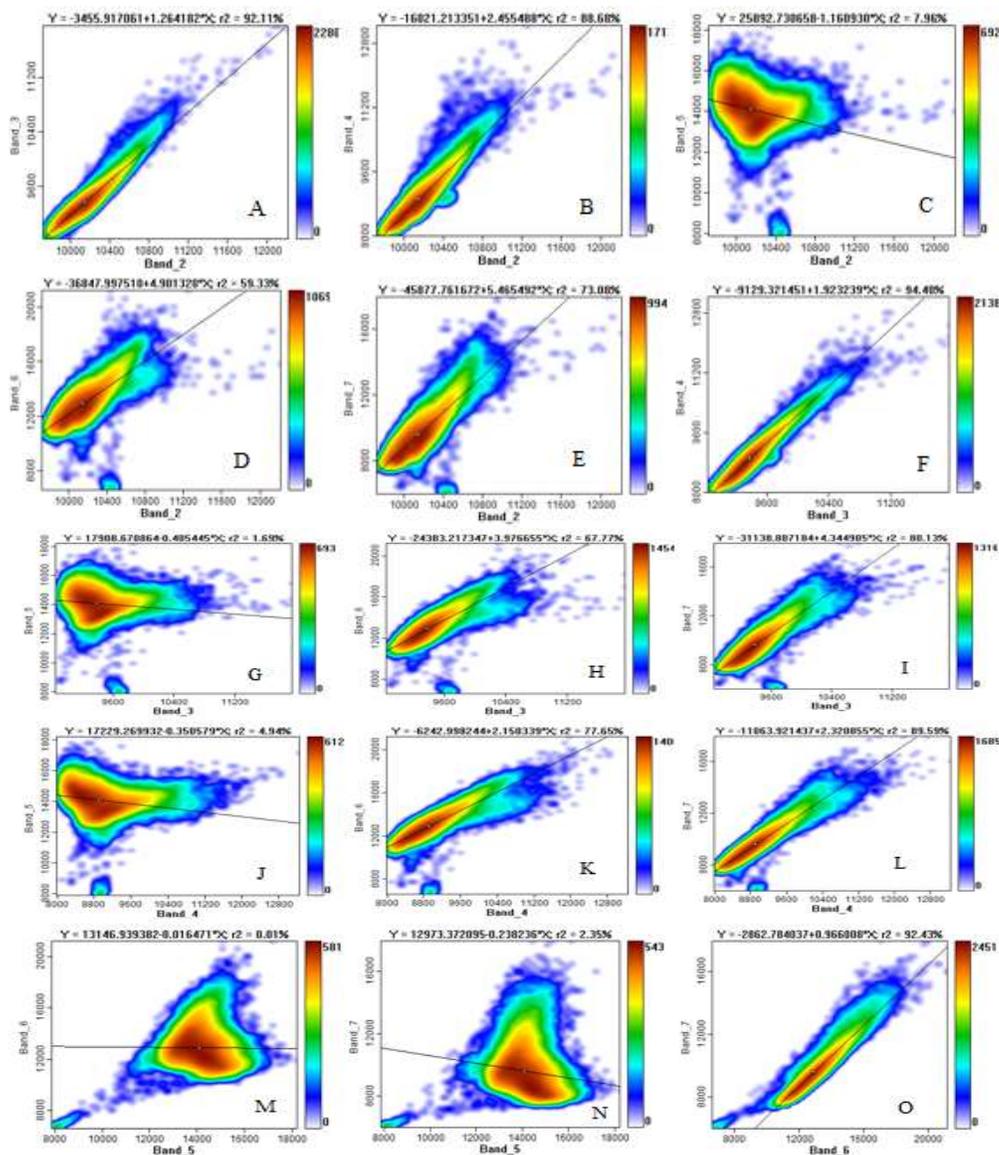
The result of the correlation matrix of the raw bands (2 to 7) of Landsat 8 (OLI/TIR) showed that the highest correlations are between bands 3 and 4, and bands 2 and 3 with correlations of 0.97 and 0.98 respectively (Table 2). Some of the bands have negative correlations such as bands 2 and 5 (-0.25), bands 3 and 5 (-0.13), bands 4 and 5 (0.17) and bands 5 and 7 (-0.04). These indicate very low inter-band correlation of spectral data. When there is high inter bands correlation (Table 2), it is a pointer to the fact that, the bands contain almost the same spectral information. Therefore, the use of bands with high inter-band correlation in data processing often leads to multi-collinearity problem in data analysis. For instance, bands 2 and 3 are highly correlated, it is therefore, pertinent to pick only one of the bands for the prerequisite image analysis. The same rule applies to bands 3 and 4 with correlation of 0.98. The inter-bands correlations with coefficient of determination (R<sup>2</sup>) in Figure 2 show a more statistical representations of inter-bands relationships. Interpreting the factor-loading pattern (Table 3), of the relationships amongst bands vis-à-vis a given PC, a band is said to load heavily on a given PC if the factor loading is greater or equal to 0.50 (Alphonsus and Raji 2019). Bands 6 and 7 loaded heavily on PC1 (0.65982 and 0.68504), while PC2 is heavily loaded with Band 5 (0.95140).

**Table 2:** Correlation matrix

	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7
Band 2	1					
Band 3	<b>0.97321</b>	1				
Band 4	0.95609	<b>0.97812</b>	1			
Band 5	-0.24773	-0.13486	-0.17275	1		
Band 6	0.71139	0.74701	0.82025	0.14278	1	
Band 7	0.84718	0.87580	0.92971	-0.04247	0.94668	1

**Table 3** Relationship amongst Landsat bands expressed as loadings in Principal Component Analysis

BANDS	Principal Components (PCs)					
	PC1	PC2	PC3	PC4	PC5	PC6
Band 2	0.09298	-0.08267	0.26059	0.26003	<b>0.75536</b>	<b>0.52765</b>
Band 3	0.12880	-0.07371	0.39593	0.28863	0.30190	<b>-0.80421</b>
Band 4	0.26449	-0.15736	<b>0.52696</b>	0.46932	<b>-0.57968</b>	0.26704
Band 5	0.01274	<b>0.95140</b>	0.29973	-0.05707	-0.00831	0.03880
Band 6	<b>0.65982</b>	0.20356	<b>-0.58848</b>	0.41745	0.02956	-0.04178
Band 7	<b>0.68504</b>	-0.12793	0.24797	<b>-0.67178</b>	0.03620	0.01601



**Fig. 2** Inter-band correlation with coefficient of determination ( $R^2$ ) for (a) bands 2 and 3, (b) bands 2 and 4, (c) bands 2 and 5, (d) bands 2 and 6, (e) bands 2 and 7, (f) bands 3 and 4, (g) bands 3 and 5, (h) bands 3 and 6, (i) bands 3 and 7, (j) bands 4 and 5, (k) bands 4 and 6, (l) bands 4 and 7, (m) bands 5 and 6, (n) bands 5 and 7

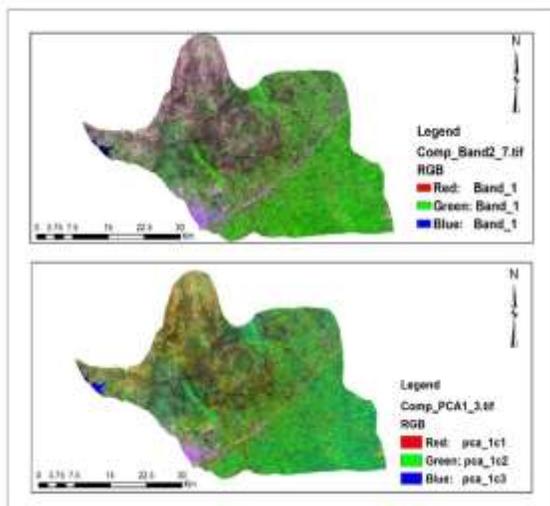
Other loadings in Table 3 are PC3 (0.52696, and -0.58848), PC4 (-0.67178), PC5 (0.75536 and -0.57968) and PC6 (0.52765 and -0.80421) respectively.

*Principal component analysis in classification:* The six Landsat raw bands were subjected to Principal Component Analysis (PCA) which resulted in a percentage of variance of the six PCA results as shown in Table 4. The results showed that PC<sub>1</sub> (81.34%), PC<sub>2</sub> (15.69) and PC<sub>3</sub> (2.3376) are the highest in terms of percentage of variance. It suffices therefore, to state that the first three PC's of the original Landsat (OLI) data described 99.37 % of the original Landsat dataset, while the last three PC's accounted for only

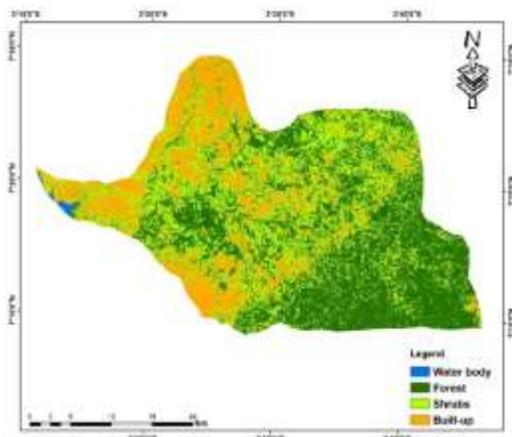
0.63% of the original dataset. In this study, raw bands of Landsat 8 (OLI/TIR) satellite image (bands 2 to 7) and PC<sub>1</sub>, PC<sub>2</sub> and PC<sub>3</sub> were composited and used for land use land cover (LULC) classification (Fig. 3). Visualizing the two composite outputs (Fig. 3), there appears to be dissimilarities between the two composite images. Supervised classification was carried out using the maximum likelihood algorithm classifier. Four thematic classes (water body, forest, shrubs and built-up) were classified in the study area (Fig. 4 and 5). Though, the main objective of this study is not premised on the LULC classes, it suffices to look at the statistical derivatives from the two LULC classifications.

**Table 4** Percent and accumulative eigenvalues

PC	Eigen Values	% of Variance
PC_1	4824198.42908	81.3392
PC_2	930726.62989	15.6927
PC_3	138642.09620	2.3376
PC_4	32538.82205	0.5486
PC_5	3886.74969	0.0655
PC_6	968.02130	0.0163

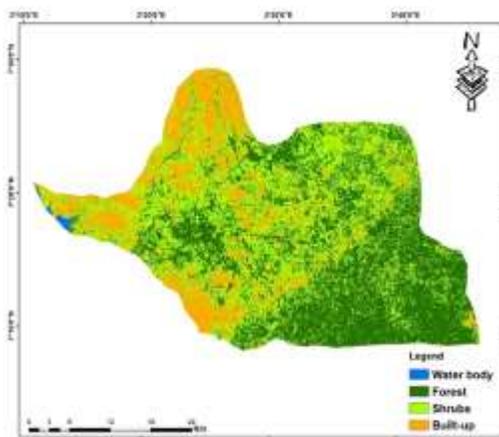


**Fig. 3** Composite images of Bands 2-7 and PC 1-3



**Fig.4** LULC classification with bands 2 to 7.

Classification results (Table 5) showed disparities between LULC classification with bands 2 to 7 composite (LULC\_Bands) and LULC classification based on PCs 1 to 3 (LULC\_PCA). For instance, forest class in LULC\_Bands occupied 41% of the total study area while the forest class in LULC\_PCA occupied 39% with a difference of 2%. Shrubs class in LULC\_Bands was 33% while LULC\_PCA remained 39%. Built-up was 26% (LULC\_Bands) and 22% (LULC\_PCA) respectively.



**Fig. 5** LULC classification with PCA dataset

**Table 5** Statistics of LULC with bands and with PCA composites.

LULC	LULC with bands		LULC PCA		Difference	
	Area (ha)	%	Area (ha)	%	Area (ha)	%
Water body	421.38	0	421.38	0	0	0
Forest	57368.96	41	57227.9	39	141.06	2
Shrubs	46067.49	33	53629.65	39	7562.16	6
Built-up	35304.94	26	27883.79	22	7421.15	4
<b>Total</b>	<b>139162.77</b>	<b>100</b>	<b>139162.77</b>	<b>100</b>	<b>15124.37</b>	<b>12</b>

*Accuracy Assessment:* As a quantitative method of characterizing image classification precision, a confusion matrix (or error matrix) is generally used. It is a table showing correspondence between the consequence of the classification and the reference image. In error matrix table, cells indicate the amount of correlation between ground truth image and classified image. Diagonal elements in the matrix give the correctly identified pixels. Kappa coefficient is a very important parameter and its value ranges between 0 and 1. Results of Error Matrices for LULC\_Bands and LULC\_PCA are shown in Tables 7 and 8. These tables show correlation matrices of reference data (Google Earth Image) of the study area and the Landsat satellite image of the same study area. The cells in the matrix tables indicate the amount of correlation between reference image and classified image with specific reference to the LULC classes. A total of 508 random points were chosen for the accuracy assessment of this study. The error matrices for the two classified images were calculated and their performances were statistically compared as shown in Tables 6 and 7.

**Table 6** Error matrix for LULC\_Bands

Land cover	Water body	Forest	Shrubs	Built-up	Row Total
Water body	9	1	2	0	12
Forest	1	243	4	6	254
Shrubs	0	9	76	7	92
Built-up	0	7	1	147	155
Column Total	10	260	83	155	508

Comparing the User’s Accuracy of LULC\_Bands and LULC\_PCA, Water body (0.75, 0.77), Forest (0.96, 0.96), Shrubs (0.83, 0.88) and Built-up (0.95, 0.96) showed some differences in statistics. Conversely, The Producer’s Accuracy had Water body (0.90, 1.00), Forest (0.93, 0.95), Shrubs (0.92, 0.93) and Built-up (0.92, 0.93) respectively. The Kappa Coefficients of both LULC\_Bands and LULC\_PCA are also comparable. While that of LULC\_Bands was 0.880113 that of LULC\_PCA remained 0.908246.

**Table 7** Error matrix for LULC\_PCA

Land cover	Water body	Forest	Shrubs	Built-up	Row Total
Water body	10	1	2	0	13
Forest	0	248	3	6	257
Shrubs	0	6	77	5	88
Built-up	0	5	1	144	508
Column Total	10	260	83	155	508

**Table 8** Performance Analysis of the accuracy assessment

LULC	User’s Accuracy		Producer’s Accuracy		Kappa Coefficient	
	LULC_Bands	LULC_PCA	LULC_Bands	LULC_PCA	LULC_Bands	LULC_PCA
Water	0.75	0.77	0.90	1.00	0.880113	
Forest	0.96	0.96	0.93	0.95	0.908246	
Shrubs	0.83	0.88	0.92	0.93		
Built-up	0.95	0.96	0.92	0.93		

Though, the observable accuracy differential could be traceable to so many underpinning factors, it suffices to infer, that the PCA operations carried out on the dataset may have contributed to the recorded higher Kappa Coefficient results (Table 8)

**Conclusion:** This research showed that the PCA approach was a useful image pre-processing technique to diminish the dimensionality of data for the study area. The first three PCs derived from the raw bands contain most of the information of the original data. Using the first three PC results in better classifications than the original dataset. The accuracy assessments of the two LULC types added credence to the importance of image pre-processing using PCA.

**Abbreviations:** PCA: Principal Component Analysis; LULC: Land use land cover; ML: Maximum Likelihood; OLI: Operational land imager; TIR: Thermal Infra-red; NPC: National Population Commission.

**Acknowledgements:** The authors appreciate Forestry Research Institute of Nigeria for providing the enabling environment for this research work. There was no financial support from any individual or corporate organization for this research and for the publication of this study.

**REFERENCES**

Adedeji, OH; Olayinka, OO; Tope-Ajayi, OO; Azeez, OI (2020). Land use/Land cover changes due to Quarrying in Odeda Local Government Area of Ogun State, Nigeria: An Assessment and Implication for Rural Livelihood. *J. Met & Clim. Sci.* 18(1):28-40.

Alphonsus, C; Raji, AO (2019). Application of Principal Component Analysis (PCA) for correcting multicollinearity and dimension reduction of morphological parameters in Bunaji Cows. *Nigerian J. Anim. Sci.* 21(2):1-8

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

Bair, E; Trevor, H; Paul, D; Robert, T (2006). Prediction by Supervised Principal Components. *Journal of the American Statistical Association.* 473(19):119-137.

Congalton, RG (1991). A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sens. Environ.* 37:5–46

David Gómez-Palacios (2017). Flood mapping through principal component analysis of multitemporal satellite imagery considering the alteration of water spectral properties due to turbidity conditions. In *Geomatics, Natural Hazards and Risk* 8(2):607–623.

Dwivedi, A; Tolambiya, A; Kandula, P; Subhash, N; Bose, C; Kumar, A; Kaira, PK (2006). Color Image Compression Using 2- Dimensional Principal Component Analysis (2DPCA) | *Proc. of ASID* p.488-491.

Estornell, J; Mart, GM; Sebast,i MT; Mengual J (2013). Principal component analysis applied to remote sensing. *MSEL.* 6:83–89.

Gao, J; Chai, S; Zhang, B; Xia, Y (2019). Research on network intrusion detection based on incremental extreme learning machine and adaptive principal component analysis. *Energies* 12:1223.

- Gasmi, A; Gomez, C; Zouari, H; Masse, A; Ducrot, D (2016). PCA and SVM As Geo Computational Methods for Geological Mapping in The Southern of Tunisia, Using Aster Remote Sensing Dataset. *Arab. J. Geosci.* 9:753.
- Geng, D; Zhang, H; Wu, H (2020). Short-term wind speed prediction based on principal component analysis and lstm. *Appl. Sci.* 10:4416.
- Imran, S; Bajwa, S; Irfan, H (2005). PCA based Image Classification of Single-layered Cloud Types. *Journal of Market Forces* 1(2):3-13
- Jeeva, D; Lakshmi, S. Naraana, Reddy (2016). Land Cover Classification based on NDVI using LANDSAT8 Time Series: A Case Study Tirupati Region, in International Conference on Communication and Signal Processing p.1332-1335.
- Klainbaum, DG; Kupper, LL; Muller, KE (1998). *Applied Regression Analysis and Multivariable Methods*. 3rd Edition, Colle Pacific Grove, CA. p. 254.
- Lan, L; Li, CF; Lei, YM; Yin, JY; Zhao, JJ (2017). Feature Extraction for Hyperspectral Remote Sensing image using weighted PCS-ICA. *Arab. J. Geosci.* 10:307.
- Li, H; Zhang, Z; Yin, X (2020). A novel probabilistic power flow algorithm based on principal component analysis and high-dimensional model representation techniques. *Energies* 13:3520
- Marchetti, MO; Roy, M; Abdat, F; Branchu, P; Bourson, P; Jobard, C; Durmont, JF; Casteran, G (2020). Chemometrics-assisted monitoring in raman spectroscopy for the biodegradation process of an aqueous polyfluoroalkyl ether from a fire-fighting foam in an environmental matrix. *Environments* 7:4.
- Masoumeh, KD; Amir, EB (2016). Use of Principal Component Analysis in Accuracy of Classification Maps (Case Study: North of Iran. *Research Journal of Forestry* 10(1):23-29.
- Meera, GG; Parthiban, S; Nagaraj, T; Christy, A (2015). NDVI: Vegetation change detection using remote sensing and GIS – A case study of Vellore District,” 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015), *Science Direct, Procedia Computer Science* 57:1199–1210
- Mofarreh, MD (2015). Color Image Compression using PCA. *IJCA* (0975-8887) 111(5):16-19.
- Munyati, C (2004). Use of Principal Component Analysis (PCA) of remote sensing images in wetland change detection on the Kafue Flats, Zambia. *Geocarto Int.* 19:11-22.
- Mussie, O (2011). Bias as Land Cover Changes estimates due to mis-registration. *International Journal of Remote Sensing* 21:3553-3560.
- Nagamani, K; Jayakumar, L; Yasodharan, S; Sriganesh, J (2015). Study on Error Matrix Analysis of Classified Remote Sensed Data for Pondicherry Coast. *J. Adv. Res. GeoSci. Rem. Sens.* 2:3-4
- National Population Commission (NPC) (2006). Nigerian National Census: Population Distribution by Sex, State, LGAs and Senatorial District: 2006 Census Priority Tables (Vol. 3). <http://www.population.gov.ng/index.php/publication/140-popn-distri-by-sex-state-jgas-and-senatorial-distri-2006>
- Ross, S; Lunetta, JG; Lyon (2004). Remote sensing and GIS accuracy assessment. CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.
- Schwartz, C; Ramos, LP; Duarte, LT; Pinho, MDS; Pettersson, MI; Vu, VT; Machado, R (2020). Change detection in UWB SAR images based on robust principal component analysis. *Remote Sens.* 12:1916.
- Venkata, M; Dasu, PVN; Reddy, S; Chandra, MR (2019). Classification of Landsat-8 Imagery Based on PCA And NDVI Methods *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 8(10):2278-3075
- Zhang, Y; Mishra, RK (2012). A review and comparison of commercially available pan sharpening techniques for high resolution satellite image fusion. In *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE International 1:182-185.
- Zhao, X; Wang, W; Chu, X; Li, C; Kimuli, D (2017). Early detection of aspergillus parasiticus infection in maize kernels using near-infrared hyperspectral imaging and multivariate data analysis. *Appl. Sci.* 7:90.
- Zhu, Z.; Yang, SV; Stehman, RL; Czaplewski (2000). Accuracy assessment for the U. S. Geological Survey regional land-cover mapping program: New York and New Jersey region, *Photogram. Eng. Remote Sens.* 66:1425–1435.