Image Classification and Accuracy Assessment Using the Confusion Matrix, Contingency Matrix, and Kappa Coefficient

F. F. Howard, C. B. Boye, I. Yakubu, J. S. Y. Kuma

***Abstract*—**One of the ways that could be used for the production of land use and land cover maps by a procedure known as image classification is the use of the remote sensing technique. Numerous elements ought to be taken into consideration, including the availability of highly satisfactory Landsat imagery, secondary data and a precise classification process. The goal of this study was to classify and map the land use and land cover of the study area using remote sensing and Geospatial Information System (GIS) analysis. The classification was done using Landsat 8 satellite images acquired in December 2020 covering the study area. The Landsat image was downloaded from the USGS. The Landsat image with 30 m resolution was geo-referenced to the WGS_84 datum and Universal Transverse Mercator (UTM) Zone 30N coordinate projection system. A radiometric correction was applied to the image to reduce the noise in the image. This study consists of two sections: the Land Use/Land Cover (LULC) and Accuracy Assessments using the confusion and contingency matrix and the Kappa coefficient. The LULC classifications were vegetation (agriculture) (67.87%), water bodies (0.01%), mining areas (5.24%), forest (26.02%), and settlement (0.88%). The overall accuracy of 97.87% and the kappa coefficient (K) of 97.3% were obtained for the confusion matrix. While an overall accuracy of 95.7% and a Kappa coefficient of 0.947 were obtained for the contingency matrix, the kappa coefficients were rated as substantial; hence, the classified image is fit for further research.

***Keywords*—**Confusion Matrix, contingency matrix, kappa coefficient, land used/ land cover, accuracy assessment.

I. Introduction

Soil offers the foundation for plant growth, a medium for water storage, and a habitat for organisms [6]. Soil constitutes an important resource and an essential part of the environment and ecosystem, from which most of the worldwide food requirements are met. The soil is one of the most valuable sources for the sustenance of life. The world's ecosystems are impacted in a far-reaching way due to soil degradation [10].

Scientifically, the main cause of soil degradation is fluvial soil erosion [1]. The activities of illegal small-scale miners, popularly known as galamsey, have contributed immensely to land degradation in many mining communities in Ghana, especially in Wassa Amenfi East District. The operations of these illegal miners are not regulated, and education has not been a prerequisite for their work [9]. These have led to the destruction of large tracts of land, including forest reserves and other fertile farmlands, changing the face of the landscapes and their scenery.

This research used a supervised classification technique to identify and quantify the study area's LULC and assess the accuracy of the classification using the error matrix approach. LULC classification performs an important function in planning an improvement scheme for an area or nation [5].

The error matrix is a standard way of analyzing the accuracy of remote sensing image classification, such as land use and land cover changes [2]-[4], [7], [11]. The significance of the classified map is a function of the accuracy of the classification. Hence, accuracy assessment is an important step in analyzing remote sensing data. This research used an error matrix accuracy assessment for the classified map of Wassa Amenfi East District to evaluate the accuracy of the remotely sensed image classification.

II. Study Area

The Wassa Amenfi East District is located in the Western Region of Ghana. The district can be found in the middle of the region. It lies between latitudes 5° 30′ N and 6° 15′ N and longitudes 1° 45′ W and 2° 11′ W. It is bounded to the west by Wassa Amenfi Central District, to the East by Upper Denkyira East District, to the south by Preastea Huni Valley District Assembly, and to the north by Upper Denkyira West District, as shown in Fig. 1. The capital, Wassa Akropong, is 180 km away from the regional capital, Sekondi-Takoradi, and 136 km from Kumasi by road.

III. Materials and Methods Used

A. Materials Used

TABLE I
Acquisition of Data for the Study

SN	Software	Usage
1	Erdas Imagine, 2014	For preprocessing of the satellite data (Landsat 8) for 2020. Radiometric and geometric correction of the data was carried out. It was also for the layer stacking of the multispectral bands of the processed satellite data.
2	ArcMap, 10.8	For displaying and processing supervised classification, LULC.
3	Google Earth Pro	For identifying the various land cover classes during the supervised classification
4	Microsoft Excel, 2016	For accurate calculation of the area and percentage of the various land uses and formatting all data.

F. F. Howard, Postgraduate Candidate, C. B Boye, *Yakubu I., are with Geomatic Engineering, and J. S. Y Kuma is with Geological Engineering Department, University of Mines and Technology (UMaT), Tarkwa, Ghana (*e-mail: yissaka@umat.edu.gh).

B. Method Used

The satellite imagery and Google Earth were geo-referenced in the WGS_84 datum UTM Zone 30N coordinate projection system.

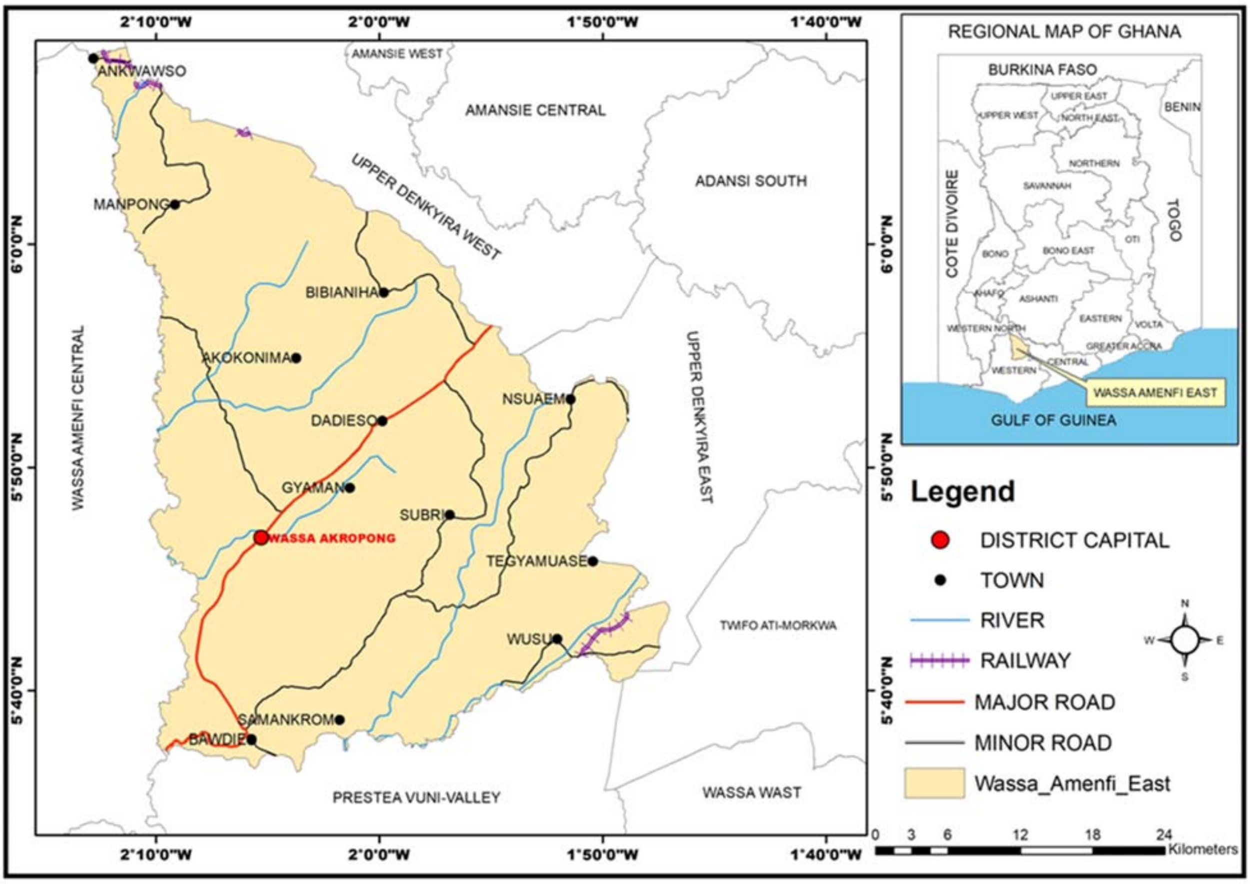

Fig. 1 Map of Wassa Amenfi East District

Supervised classification was performed according to Eastman, 2003, where the user develops the spectral classes, such as waterbody and forest, and then the software assigns each pixel in the image to the cover type to which its sign is most similar. Supervised classification is the method most commonly used for quantitative analyses of remote sensing image data [12]. The supervised classification was applied after the defined Area of Interest (AOI), which is the training classes. The AOI of a training area for a particular class was selected evenly to represent the class in the study area. The supervised classification was applied to delineate the trained classes, as shown in Figs. 2 and 3 for the methodology and the LULC map of the study area.

Accuracy Assessment in Image Classification

A confusion matrix is used to evaluate the performance of a classifier, and it tells how accurate a classifier is in making predictions about the classification. A contingency table is used to evaluate association rules.

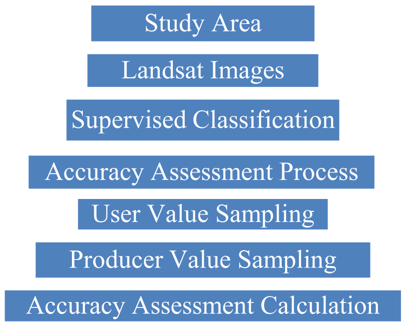

Fig. 2 Schematic workflow for accuracy assessment

Accuracy Assessment Using the Confusion Matrix (Error Matrix)

The accuracy assessment is the final stage of satellite image classification [8]. The accuracy assessment used in evaluating the classification was the error matrix [4]. The columns of the error matrix indicate the classes the pixels are in the ground truth, and the rows indicate the classes the image pixels have been assigned to in the image. The diagonal of the error matrix indicates the pixels that have been classified correctly.

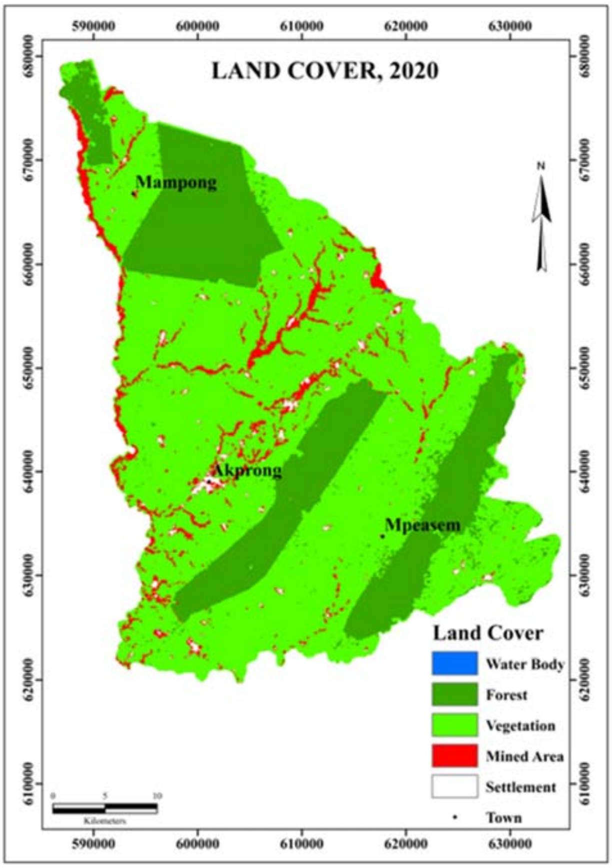

Fig. 3 LULC Map of the Study Area

In this study, the number of ground-truthing points taken to represent the LULC classes was 47, and their IDs were arranged in ascending order.

TABLE II
LAND USE/LAND COVER CLASSES

Land Use/Land Cover Classes	Assigned ID
Waterbody	1
Settlement	2
Forest	3
Vegetation	4
Mined-out Area	5

It was ensured that the ground truth points are representative of the LULC classes through an examination of Google Earth and drone images, as shown in Figs. 4-9. Shapefile points were converted to a KML file. The KML file was then opened in Google Earth. The user values (ground-truthing points) were checked to see whether they were correct or wrong on Google Earth, and the producer value was numbered, as illustrated in Table III. If user value 1 in a waterbody is correct in the Google Earth image, then the producer value will also be the same as the user value. If user value 5 for mined-out area is wrong in Google Earth, then producer value will be "vegetation" in Google Earth.

Fig. 4 Waterbody correctly classified on Google Earth

Open Science Index, Computer and Information Engineering Vol:17, No:9, 2023 publications.waset.org/10013249.pdf

Fig 5 Settlement correctly classified on Google Earth

Fig. 6 Forest correctly classified on Google Earth

Fig. 7 Vegetation correctly classified on Google Earth

Fig. 8 Mined area correctly classified on Google Earth

Fig. 9 Vegetation wrongly classified as mined area on Google Earth

Accuracy Assessment Using Contingency Matrix

The contingency matrix is a qualitative analysis using the GNSS receiver location of the Ground Truth Points of the classified images on the satellite image to know which pixel value they are associated with.

Overall Accuracy = sum of diagonal entries/total number of samples

Adjusted Accuracy = calculation of the Kappa coefficient

Mapping-category-level accuracy describes the accuracy of the individual mapping category that is present on a map.

Omission error = non-diagonal column total/column total

Commission error = non-diagonal row total/row total

Producer's accuracy = 100% minus omission error

User's accuracy = 100% - commission error

IV. Results and Discussion

A. Accuracy Assessment Using Confusion Matrix (Error Matrix)

The error matrix is computed by comparing the results obtained by the image interpretation as a column with those obtained by the reinterpretation as a row. 47 sample locations were taken randomly for the study area from 5 classes. 10 samples were taken from each of the classes, with the exception of the waterbody, which had 7 samples (Table III).

TABLE III
Error Matrix for Accuracy Assessment

	Water body	Settlement	Forest	Vegetation	Mined area	Total (Use)
Waterbody	7	0	0	0	0	7
Settlement	0	10	0	0	0	10
Forest	0	0	10	0	0	10
Vegetation	0	0	0	10	0	10
Mined-out area	0	0	0	1	9	10
Total (Producer)	7	10	10	11	9	47

Accuracy Assessment Formula

Overall Accuracy =

$$\frac{Total\ No.of\ correctly\ classified\ Pixels\ (diagonal)}{Total\ No.of\ Reference\ Pixels} \quad (1)$$

Diagonal of the Error Matrix = 7 + 10 + 10 + 10 + 9= 46
Overall Accuracy = (÷)x 100 = 97.87%

User Accuracy (UA)

$$\frac{No.of\ correctly\ classified\ pixels\ in\ each\ category}{Total\ No.of\ classified\ pixels\ in\ that\ category\ (the\ Row\ Total)} \times 100 \quad (2)$$

User Accuracy Calculation

- Waterbody = (7/7) x 100 = 100%
- Settlement = (10/10) x 100 = 100%
- Forest = (10/10) x 100 = 100%
- Vegetation = (10/10) x 100 = 100%
- Mined area = (9/10) x 100 = 90.0%

Producer Accuracy (PA)

$$\frac{No.of\ correctly\ classified\ Pixels\ in\ each\ category}{Total\ No.of\ Reference\ Pixels\ in\ each\ category\ (\ the\ column\ Total)} \times 100 \quad (3)$$

Producer Accuracy Calculation

- Waterbody = (7/7) x 100 = 100%
- Settlement = (10/10) x 100 = 100%
- Forest = (10/10) x 100 = 100%
- Vegetation = (10/11) x 100 = 90.9%
- Mined area = (9/9) x 100 = 100%

Tables IV-IX indicate the various computations of the LULC classes.

TABLE IV
OVERALL ACCURACY

Total No. of Corrected Classified Pixel (Diagonal)	45.0
Total Sample Size	47.0
Overall Accuracy (%)	95.7

TABLE V
USER ACCURACY

Land Cover	Value (%)
Waterbody	100.0
Forest	100.0
Vegetation	90.9
Mined Out	88.9
Settlements	100.0

TABLE VI
PRODUCER ACCURACY

Land Cover	Value (%)
Water Body	100.0
Forest	100.0
Vegetation	90.9
Mined Out	88.9
Settlements	100.0

Kappa Coefficient (T)

$$= \frac{(TS\ x\ TCS) - \Sigma(Column\ total\ x\ Row\ Total)}{TS^2 - \Sigma(Column\ Total\ x\ Row\ Total)}\ x\ 100 \qquad (4)$$

where TS = Total sample = 47; TCS = Total corrected sample (diagonal) = 46.

$$T = \frac{(47\ x\ 46) - [(7*7) + (10*10) + (10*10) + (11*10) + (9*10)]}{47^2 - [(7*7) + (10*10) + (10*10) + (11*10) + (9*10)]}\ x\ 100$$

$$T = \frac{2162 - 449}{2209 - 449}$$

$$T = 97.3\%$$

B. Accuracy Assessment Using Contingency Matrix

The contingency matrix is computed by comparing the locations of the ground truth obtained from the classified images to the satellite image to determine its corresponding pixel values. 47 ground truth sample locations were taken randomly within the 5 classes for the study. 10 samples were taken from each of the classes, with the exception of the waterbody, which had 7 samples.

Equations (1)-(4) give the various accuracy assessment calculation formulae.

The overall accuracy obtained for the confusion matrix was 97.87% and 95.7% for the contingency matrix. The broad range of accuracy indicates good and acceptable results for the supervised classification of the LULC over the study area. The measure of the producer's accuracy (sensitivity) reflects the accuracy of the prediction of the particular category, and the user's accuracy reflects the reliability of the classification to the user. The user's accuracy is the more relevant measure of the classification's actual utility in the field. Most of the LULC classifications were found to be reliable, with 100% user accuracy.

TABLE VII
KAPPA COEFFICIENT

(TSS * TCS) - SUM (Column Total * Row Total)/TSS^2 - SUM (Column Total * Row Total)*100	
Where:	
TSS = Total Sample Size	47
TCS = Total Correctly Sample	45
SUM (Column Total * Row Total)	451
(TSS * TCS)	2115
TSS^2	2209
Kappa Coefficient	94.7

TABLE VIII
COMMISSION ERROR

Commission Matrix (%)	
Land_Cover	Value (%)
Waterbody	0.0
Forest	0.0
Vegetation	9.1
Mined Area	11.1
Settlements	0.0

TABLE IX
OMISSION ERROR

Omission Matrix (%)	
Land Cover	Value (%)
Waterbody	0.0
Forest	0.0
Vegetation	9.1
Mined Out	11.1
Settlements	0.0

The commission errors reflect the points that are included in one category when they really do not belong to that category. For instance, in the case of vegetation, one point was misclassified and does not fall under this category, as shown in Fig. 8. Equally, the omission errors reflect the number of points that are not included in the category while they really belong to it. The omission error in the case of mined areas, a point that actually belongs to this category but was not categorized in that class, there was no misclassification of other classes. An overall Kappa coefficient of 0.973 (97.3%) and 94.7 (94.7%) were obtained for the confusion matrix and contingency matrix, respectively. Apart from overall classification accuracy, the above-individualized parameters give a classifier a more detailed description of the model performance of the particular class or category in his field of interest or study.

V. CONCLUSIONS

The images were classified into five classes; percentages for their land sizes were: water body (0.01%), settlement (0.88%), forest (26.02%), vegetation (67.86%), and mined area (5.24%).

Vegetation was the dominant type of land use classified, covering about 67.86% of the total study area.

Individual accuracy assessment parameters are useful to assess the model's performance in respect of a particular category or class of specific interest for the study. In the study, the accuracy assessments performed were the confusion and contingency matrix and kappa coefficient. The confusion matrix had an overall classification accuracy of 97.87% and a kappa coefficient of 0.973. The contingency matrix's overall accuracy was 95.7%, and the Kappa coefficient was 0.947. In comparison, the confusion matrix gave better accuracy to the classified image. The accuracy assessments were all found to be acceptable for the assessment of land use/cover classifications.

REFERENCES

[1] Angima, S. D., Stott, D. E., O'neill, M. K., Ong, C. K., & Weesies, G. A. (2003). Soil erosion prediction using RUSLE for central Kenyan highland conditions. Agriculture, ecosystems & environment, 97(1-3), pp. 295-308.

[2] Congalton RG. (2004), "Putting the map back in map accuracy assessment. In: Geospatial Data

[3] Congalton RG, Green K. (1999), "Assessing the accuracy of remotely sensed data: Principles and Practices". Lewis Publisher, Boca Raton,

[4] Congalton RG. (1991) "A review of assessing the accuracy of classifications of remotely sensed Data". Remote Sensing of Environment; 37(1): pp. 35–46.

[5] Disperati, L., Gonario S. and Virdis, P. (2015) "Assessment of Land-Use and Land-Cover Changes from 1965 to 2014 in Tam Giang-Cau Hai Lagoon, Central Vietnam". Applied Geography, 58, pp. 48-64.

[6] Dominati, E., Patterson, M. and Mackay, A. (2010) A Framework for Classifying and Quantifying the Natural Capital and Ecosystem Services of Soils. Ecological Economics, 69, pp. 1858-1868.

[7] Hammond TO, Verbyla DL (1996). Optimistic bias in classification accuracy assessment. International Journal of Remote Sensing; 17(6): pp. 1261-66.

[8] Manisha B. G Chitra & N Umrikar (2012). "Image Classification Tool For Land Use/ Land Cover Analysis": A Comparative Study of Maximum Likelihood and Minimum Distance Method. Int. J. Geo. earth. environ. , 6, pp. 189-96.

[9] Kuma J.S, Jerome A. Yendaw (2010), InternationalJournal of Geosciences, 2010, 1, 113-120 doi:10.4236/ijg.2010.13015 Published Online November, 2010 (http://www.SciRP.org/journal/ijg).

[10] Pouyat, R., Groffman, P., Yesilonis, I., & Hernandez, L. (2002). Soil carbon pools and fluxes in urban ecosystems. Environmental pollution, 116, pp. S107-S118.

[11] Story M, Congalton RG (1986). "Accuracy Assessment": A user's perspective. Protogrammetric Engineering and Remote Sensing; 52(3): 397-99.

[12] Richards, J., and Jia, X. (2006) Remote Sensing Digital Image Analysis: An Introduction. Springer, Berlin.