International Journal of Engineering Works
Kambohwell Publisher Enterprises
ISSN: 2409-2770 Vol. 4, Issue 1, PP. 10-20, January 2017
www.kwpublisher.com

Comparative Performance of Multi-Source Reference Data to Assess the Accuracy of Classified Remotely Sensed Imagery: Example of Landsat 8 OLI Across Kigali City-Rwanda 2015

Felicien Nkomeje

***Abstract*—**Accuracy assessment of remote sensed classified images is considered the backbone of remote sensing image processing to be considered credible. However, reference data to perform this task is also a considerable challenge for the remote sensing analyst. This study was carried out over Kigali city using Landsat remotely sensed imagery acquired on July 15, 2015, to compare multi-sourced reference data performance to assess the accuracy of classified Landsat remote sensed imagery. To achieve this objective, GeoEye-1, WorldView-2, Google earth high-resolution image, and GIS layers have been used to verify the accuracy of remote-sensed data classification. In this study, we applied different reference data sources to Landsat 2015 classified images to assess the accuracy. Therefore, results from GEOEYE-1 image as reference data source displayed the total accuracy and kappa coefficient of 98.5% and 0.98 respectively. WorldView-2 MS Image revealed 97.25% of total accuracy and a 0.96 Kappa coefficient agreement.

High-resolution rectified images generated using El-Shayal Smart GIS Editor also show its capabilities to assess the accuracy of Landsat remote sensed data whose results were 94% and 0.92%, respectively, for overall accuracy and total Kappa statistics. Furthermore, the remote sensing analyst should not worry about where or how to find reference data to assess image classification so long as they possess GIS shape files. GIS shape files provide good results where the overall accuracy was 92% and a Kappa coefficient of 0.90. Moreover, GIS shape files results showed a slightly lower accuracy because of data properties; it is recommended to check projection before using any spatial data. This paper strongly focused on soft features during ground reference data collection. Test data from GEOEYE-1 images have shown the best thematic accuracy after being overlaid with Kigali 2015 thematic map. All of the referenced data sources, in general, showed the ability to assess remote sensed classified map in the range of 90% to 98.5% for both total accuracies of the map and kappa accuracy.

Keywords: Remotely sensed data, Multi-source reference data, Thematic accuracy assessment, El-Shayal Smart GIS Editor, Kigali, Rwanda

Felicien NKOMEJE: School of Earth Science and Engineering , Hohai University, fnkomeje@gmail.com , No. 1 Xikang Road, Nanjing 210098, P. R. China, Tel: (+86) 18305160547

I. INTRODUCTION

Many significant journals report on remote sensing classification [1-3], reference data [4, 5], and accuracy assessment [6-9] of remotely sensed digital image classification, and this research cannot embrace their totality. However, it will briefly provide important aspects that focus mainly on accuracy assessment and reference data sources [10-13]. Many environmental monitoring studies necessitate the area coverage so all of these will be realized using remote sensors that acquire multiple requirements including different resolutions, band designation and others special specifications [14-16]. Classified digital image validation is a standard component of any land use and land cover study project [17]. Maps from remotely sensed data must be validated to ensure usefulness. Knowing the accuracy of the map is vital to any decision making performed using that map and its related results. The three basic components of an accuracy assessment are sampling design, response design, and estimation with analysis procedures [2, 18-20].

Hence, without reference data, the accuracy cannot be ensured since it is a confrontation of the test pixels (reference data) and classification map. The process of assessing the map accuracy is time-consuming and expensive. It is very important that the procedure is well thought out and carefully planned to be as efficient as possible. [2, 4]. The accuracy assessment of spatial data determines the information value of the resulting data to a user [6, 7, 21]. Usefulness of geospatial data is only possible if the quality of the data is recognized. Furthermore, the use of different types of geospatial data is not reliable if the data quality is not verified. Therefore, the accuracy must be calculated based on sample size and it is very important to have the significant test size. Nevertheless, the use of an excessively large sample may lead to the conclusion that any non-zero difference observed is statistically significant. Conversely, the use of a sample size that is too small may result in a failure to detect a difference that may actually be large and important [22-24].

It is imperative to judge if the classification results coincide with the nature attributes in order to be used with confidence [23]. Moreover, this is critical to end-users of the data who may need valuable metadata that will be necessary during the exchange of more standard digital spatial between them [7]. Ideally, several measures of accuracy assessments should be performed and included as documentation with the classification. This process of determining image classification accuracy resamples classified imagery against through from the ground in in field samples often obtained with a Global Positioning System [25] and high-resolution images which are time-consuming and costly or unavailable. Time, and funding are habitually a big challenge to the amount of data that can be collected since we need in situ measurements. Therefore, remote sensing analysts are obliged to proceed with high-resolution satellite imagery as a substitute which is sometimes unavailable, protected by producers or companies, or simply too hard to afford. Conversely, depending on the intended use of the thematic map, some level of accuracy assessment may be performed using the same original image used for classification. As a result, quantification of this phenomenon is rather complicated and sometimes confusing. Nonetheless, all reference data sources are not accurate on the same level and are not valid for all situations because some require great attention to be sure that the reference data from that source are accurate and useful [17].

Some researchers are still arguing on how much accuracy is really needed to be qualified as credible tet there is no single answer to this question [2, 10, 19]. The situation and decisions made at National or Regional level for land use are not the same as those needed for zoning or very local boards. Moreover, the present acceptance of 80% accuracy is resulting in poorer data provided to decision makers and this is being improved incrementally with remote sensing technology development and high-resolution image production [19]. Furthermore, remote sensing analysts need to investigate LUC of the area. They need alternative data sources available with a simple inspection to validate thematic maps produced. Nowadays, researchers are using different sources to get reference data during accuracy assessments of classified images. The most used reference data sources are GPS data from the field and high-resolution images[26]. Thus, the total accuracy and kappa coefficient have become a standard means of assessment of image classification accuracy [27]. These errors and accuracies are from a confusion matrix built from collected data (with GPS or from high-resolution images).

It is really costly and time-consuming and some areas are not accessible. All these sources are different in terms of time-consumption, financial expenses and technique, and accuracy. A team of Egyptian engineers developed a tool to gather and manage rectified Google Earth, a large, and high-resolution satellite image [28, 29]. El-Shayal Smart GIS Editor) is free and can work directly in conjunction with Google Earth to produce images that can serve as reference data sources during image classification accuracy assessments. These images can be downloaded at a different scale to enhance the visibility according to the purpose of use[28, 29]. GIS users possess shapefile in place, and they always have challenges related to lack of accurate reference data while dealing with remotely sensed data, LUC analysis and classified image accuracy assessment [30] because of the landscape changes with time. Data fusion can be analyzed in order to assess its capabilities in thematic map accuracy assessment and reference data source improvement [31].

II. Objectives of the Study

Despite the extensive and numerous publications on this topic of accuracy assessment, the basic reference data source and statistical structure of errors and accuracy of the LUC types are not yet clear. This led us to mainly compare multi-source reference data performance to assess the accuracy of classified Landsat remote sensed data across Kigali. Beyond this main objective, we have some other specific objectives which are the following: Developing a reliable methodology to use during remotely sensed image classification and results in accuracy verification; Evaluating El--Shayal Smart GIS Editor performance in conjunction with Google Earth to provide reference data image during environmental studies; Describing GIS layers usefulness as reference data source and its accurateness coupling with GPS field collected data; and Identifying the best source of reference data for remote sensed data classified accuracy assessment.

III. Methodology

A. Study area

The study unit is superficially ordered in some areas form. The surface area of Kigali city is estimated at 731 km2 [32]. The area is divided into different zones according to the changes in land use and land cover. Among those LUC, the study states built up roads and transport facilities, agricultural land, water areas, wetlands, etc [33]. The sample will be exhaustive and cover the whole study area of Kigali according to the available data. Digital images from USGS constitute primary data, reference data sources constitute secondary data, and beyond these, we will use tertiary or ancillary data to carry out this research.

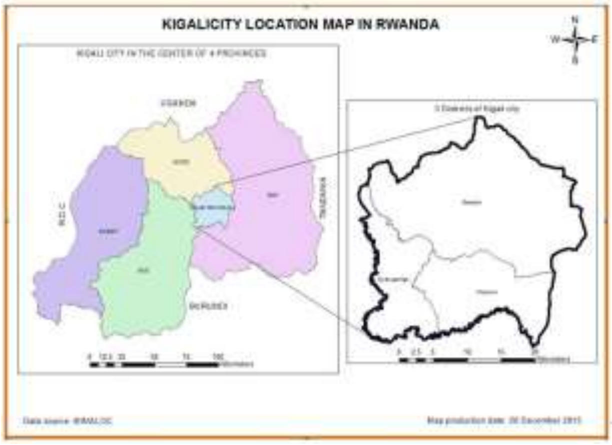

Figure 1: Kigali city location map

Source: Data from Kigali city and map produced by the Author, December, 2015

The map above shows the city of Kigali in Rwanda. Kigali is divided into 3 districts (Nyarugenge, Kicukiro, and Gasabo). Nyarugenge is the urban center. The rate of urbanism is also high in Kicukiro District, more so than in Gasabo District. A big part of Gasabo District is not developed and built-up. The city Districts' urbanism rate is in inverse proportion to their size.

ISSN: 2409-2770

B. Approach design

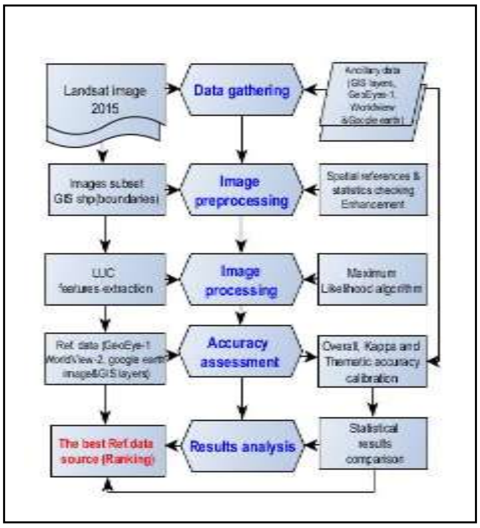

Figure 2: Research approach flowchart

C. Data source and supporting software

Landsat 8 Operational Land Imager images consist of nine spectral bands with a spatial resolution of 30 meters for Bands 1 to 7 and 9 [34].

New band 1 (ultra-blue) is useful for coastal and aerosol studies. New band 9 is useful for cirrus cloud detection.

The resolution for Band 8 (panchromatic) is 15 meters.

Thermal bands 10 and 11 are useful in providing more accurate surface temperatures and are collected at 100 meters. Approximate scene size is 170 km north-south by 183 km east-west (106 mi by 114 mi). TIRS bands are acquired at 100-meter resolution but are resampled to 30 meters in delivered data product.

In this study, we also used different datasets to assess the accuracy of image classification. Those datasets are GEOEYE-1 with 0.5m of resolution and WorldView-2 with 2m of resolution images both of which aquired in 2015; Google Earth high-resolution images downloaded from google earth in September 2016 using El Shayal Smart GIS Editor software with 3m of resolution; and GIS datasets updated in 2015 during the Kigali master plan upgrade at a different scaled levels.

Except for a very small number of Landsat TM scenes which are processed using the National Land Archive Production System (NLAPS), all Landsat standard data products are processed using the Level 1 Product Generation System (LPGS) with the following parameters applied: GeoTIFF output format, Cubic Convolution resampling method, 30-meter (TM, ETM+) and 60-meter (MSS) pixel size (reflective bands), Universal Transverse Mercator (UTM) map projection (Polar Stereographic projection for scenes with a center latitude greater than or equal to -63.0 degrees), World Geodetic System (WGS) 84 datum and MAP (North-up) image orientation.

All these sources, especially GIS datasets, had to be checked and updated using points collected from the field in order to avoid ill-fitting results from different projection problems [35, 36].

Table 1: Characteristics of original classification image and reference data source datasets

Reference year	Date of acquisition	Spacecraft ID	Sensor ID	Mission life	WRS_ P/R	Number of bands	Spatial resolution
2015	12 July 2015	Landsat_8	OLI	February 2013 up to date	172/61	11	30m 15m 100m/30m
Geometric descriptions of Landsat images used are the following: WRS_P/R: 172/61, MAP_PROJECTION = "UTM", DATUM = "WGS84", and UTM_ZONE = 36S.							

ISSN: 2409-2770

The following are Landsat 8 OLI and multi-source datasets we used to extract reference pixels that helped to evaluate remotely sensed imagery classification results (Thematic map):

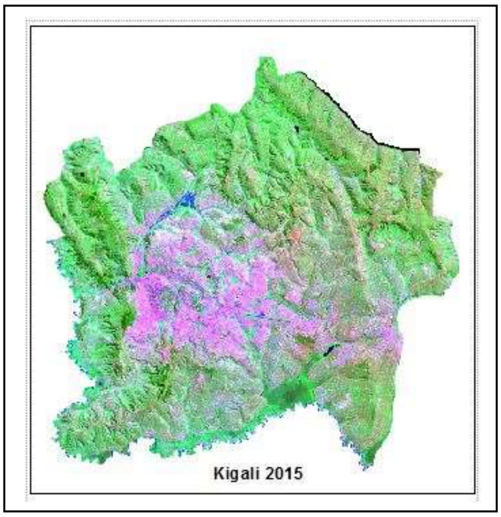

Figure 3: Landsat 8 OLI scene

Figure 5: GeoEye-1 scene

Figure 7: Google earth image

Figure 4: El-Shayal Smart GIS Editor interface

Figure 6: WorldView-2 scene

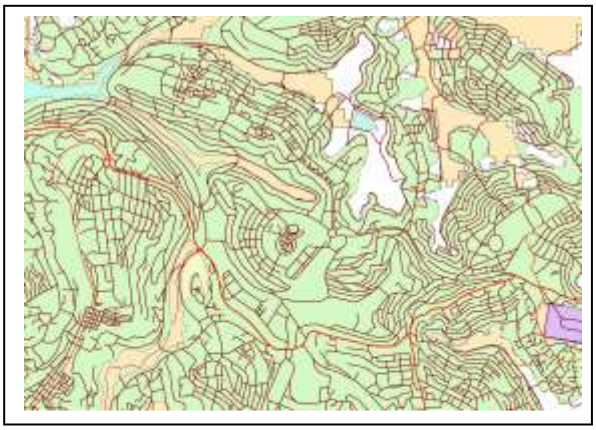

Figure8: GIS Layers

The Landsat 8 OLI image (Fig. 3) was captured in 2015; El-Shayal GIS Smart Editor coupled with Google earth (Fig. 4) was used to generate high-resolution images-3m of resolution (Fig. 7) to be used in classification assessments. Both GeoEye-1 image-0.5m of resolution (Figure 5) and Worldview images-2m of resolution (Figure 6) were used as reference data since we can clearly extract information that we need for assessing our classification reliability. GIS layers (Figure 8) include urban built-up areas, roads network layers, not-suitable for urbanization and agriculture areas, those suitable for agriculture, wetland areas, forests areas, and aquatic areas.

To execute this research the following software has been selected: ENVI classic 5.1 used to process Landsat remote sensing data, ArcGIS 10.2 to deal with geospatial analysis, Google earth coupled with El-Shayal Smart GIS Editor to generate Google earth high resolution images, Microsoft Excel 2013 to execute statistical analysis, yEd graph editor to diagram program and EndnoteX7 for bibliography management.

D. Classifier algorithm and Classification scheme

As the author masters the study area very well, the supervised classification would be more convenient to extract information from Landsat images [27, 37-39]. The land use and land cover classification will be based on pixels and the algorithm classifier has the best likelihood. Supervised classification is advantageous because it uses a relatively small number of classes to determine the appropriate land cover for each pixel. This allows for a streamlined and focused analysis. The disadvantage of supervised classification is that it requires a lot of user input prior to performing any classifications. This portion of the analysis is time-consuming, and if there are any user-induced errors, the user will have to restart the training site selection process, possibly more than once [4, 24, 40]. This classification uses the training data by means of estimating means and variances of the classes, which are used to estimate probabilities and also consider the variability of brightness values in each class as:

$$L_k(X) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_k|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(X-\mu_k)\Sigma_k^{-1}(X-\mu_k)^t\right] \tag{1}$$

Where, n: number of bands, X: image data of n bands, Lk(X) : the likelihood of X belonging to class k, μk : mean vector of class k, Σk : a variance-covariance matrix of class k, │∑k│ is a determinant of ∑k.

The amount of LULC information that may be found on the remotely sensed image depends on sensors quality and, altitude which define spatial resolution [20]. In this study, based on land use categories used by other agencies and Anderson level I LULC classification system, we defined 5 units of features to represent land use and land cover of Kigali city[20]:

1. **Urban areas or Built-up Land**: Residential, Commercial, Services, Industrial, Transportation, Communications and Utilities, Industrial and Commercial Complexes, Mixed Urban or Built-up Land, Other Urban or Built-up Land.

2. **Agricultural Land**: Cropland and Pasture, Nurseries, Other Agricultural Land, Herbaceous Rangeland and Mixed Rangeland.

3. **Forest Land**: Deciduous Forest Land, Evergreen Forest Land, and Mixed Forest Land.

4. **Wetland**: Marshland and swampland.

5. **Water areas**: Stream and Lake.

E. Accuracies, errors and Kappa agreement coefficient calculations methods

Overall accuracy also called total Accuracy is calculated by taking the Number of correct plots dividing a total number of plots. Diagonals represent sites classified correctly according to reference data and Off-diagonals are misclassified [41, 42].

$$O^c = \frac{\sum_{x=1}^{k} nii}{/T/} \tag{2}$$

Where, Oc stands for overall accuracy, k stand for classes, nii represent pixels correctly classified and /T/ the total number of pixels we are testing.

Kappa agreement coefficient, reflects the difference between the actual agreement and the agreement expected by chance. Kappa has zero value if the two nominal variables (Classification and reference pixels) are statistically independent, and value unity (1) if there is perfect agreement. Moreover, this is generally thought to be a stronger measure than simple percent agreement calculation, since Kappa takes into account the agreement occurring by chance [3, 4, 41, 43-45].

$$\widehat{K} = \frac{N\sum x_{ii} - \sum(x_{+i} \times x_{i+})}{N^2 - \sum(x_{+i} \times x_{i+})} \tag{3}$$

Where, N is the total number of samples, xii is the number of observations in row I and column i, x+i is the marginal totals of column i, xi+ the marginal totals of row i.

The aim of these calculations is to verify that a pixel is accurately classified which determines the probability that a pixel represents the class for which it has been assigned [9, 46]. Not only Overall accuracy (OA) and Kappa coefficient agreement (K) will be included but, we will also calculate Omission error (OE), Commission error (CE), Producer accuracy (PA) and User accuracy (UA) in order to evaluate the accuracy and errors occurred during classification for each class.

F. Reference data collection, Sample Size and Sampling method

Since our classification accuracy is based on site pecificity, higher resolution imagery is a suitable substitute for in situ data gathering [47]. Different sources can provide reference data to validate image classification such as photo interpretation, aerial reconnaissance with a helicopter or airplane, video, drive-by surveys, and visiting the area of interest on the ground with GNSS [48]. The sample must be collected to evaluate the accuracy of the LUC classification of the study area. Our reference data are GEOEYE-1 images (0.5m), WorldView-2 (2 m), Google earth image (3m) and GIS shapefiles that have been verified in the field throughout Kigali master plan upgrading using Garmin GPSMAP 62S Handheld Navigator. Reference data are compared to a classification map to make sure the class type of classified image matches the class type determined from reference data. Typically, GIS shape files are not used as Reference data sources, and this study will evaluate their capabilities once combined with some ground truth points for verification.

Sample size as a subset of a population must be appropriate. It has to derive any meaningful estimates from the error matrix [5]. The population size is based on pixel number of the whole classified image. In particular, very small and very big sample sizes can produce misleading results [49]. Sample sizes can be calculated using the equation of proportion described below [50].

$$n = \frac{N}{1+N(e)^2} \quad \rightarrow \text{ sample size estimation} \qquad (4)$$

Where, n is the sample size (Number of reference points/pixels), N is the whole population size (Total pixels of the map), and e is the level of precision error.

Our area has 814,016 pixels and we decided to use the formula for categorical data for an alpha level a priori at 0.05 (precision error of 5%).

Numerical application:

$$n = \frac{814016}{1+814016(0.05)^2} = 399.80 \cong 400$$

Also, the sample size can be calculated based on sampling error and fixed confidence level using the following formula:

$$N = \frac{Z^2\,(p)(q)}{E^2} \quad \rightarrow \text{Sample size estimation} \qquad (5)$$

Where, N is the sample size, P is the expected accuracy that we would like to achieve, q = 100 – p, and E stands for Allowable error. *Z* = 2 (from the standard normal deviate of 1.96 for the 95% two-sided confidence level). This is the "best guess" about the accuracy and area information that can be used for the sample size calculation. The number of reference pixels required for accuracy assessment depends on the minimum level of accuracy [51, 52]. According to this formula, the sample size will be 204 if the acceptable accuracy is 85%, q will be 5%, accepted errors is 15 % and z will be 2. The target accuracy of 0.85 or 85% is often suggested in remote sensing applications, although the use of this value is debatable [23].

$$N = \frac{2^2(85)(15)}{5^2} = 204$$

On the other hand, a general rule of thumb developed from many projects shows that sample sizes of 50 to 100 pixels for each map category are recommended, and each category can be assessed individually [53]. By combining formulas above, the rule of thumb statements and extracted information status which reveal the homogeneity and the high spatial frequency, we turned to the rule of thumb and decided to assign at least 50 points to the smallest class in proportion and 110 points to the class with big proportion even if it can slightly affect the results but can also pass as comparison study[54]. To raise the small proportion, reference points will be deducted from other classes having more than 110 points in order to increase the results' accuracy.

The sample size will be 400 pixels and all classes will share this figure according to their real size (Agriculture 110, water 50, built up 80, Forest 80 and wetland 80). The level of precision error must be fixed at low level. The high precision level depends on the small sample size. If the population is small or big, then the sample size can be reduced or partially completed [55]. This means that water class will have the fewest reference pixels because it is the one with the smallest number of classification pixels, and agriculture will have more than others because it presents a large number of classification pixels in the whole area. Moreover, minimum sample size will be at least 20 to 100 samples per strata (for use as a good estimation) [8, 56]. A stratified random method as an appropriate sampling method for accuracy assessment has been used. This is convenient for per-category or total accuracy basis assessment [57]. A stratified random sample is a multinomial sampling method, and therefore is an appropriate sampling method to be used with the Kappa Coefficient of Agreement [56].

IV. COMPARISON OF RESULTS AND DISCUSSION

The thematic map (fig.9) shows classification scheme spatial distribution of different features. From the thematic map, using reference pixels extracted from high resolution maps and GIS layers, the accuracy assessment has been performed and results are shown in the table 3, 4, 5, and 6.

A. *Features extraction map, Errors, and accuracy measurements calculations*

The following is the thematic map extracted from Landsat 8 image captured in 2015.

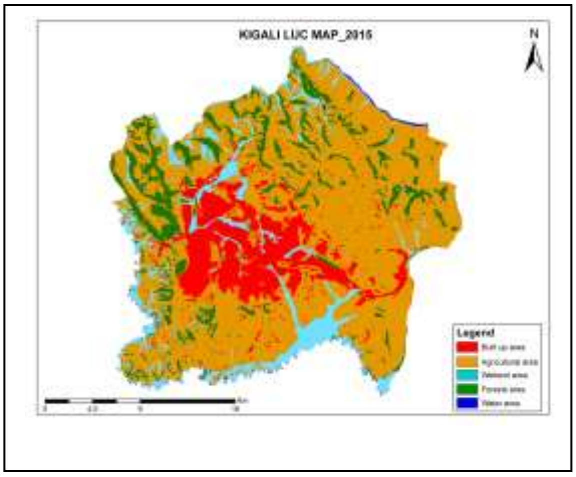

Figure 9: Kigali city classification map, 2015

The thematic Classification map was executed applying pixelbased algorithm (maximum Likelihood) and produced 5 classes. The contingency matrix (also called likelihood table) that we generated from classification and reference pixels (Reference from 4 different sources) combined allows us to calculate meaningful numbers of the data. The measurements that we computed, as shown in the table below, are the following:

GTP: Ground Truth Pixels, PA: Producer's Accuracy, UA: User's Accuracy, CE: Commission Errors, OE: Omission Errors, OA: Overall accuracy and K: Kappa coefficient

Table 2: Contingency matrix, errors, and accuracy assessment using GEOEYE-1 image as source of reference data

LUC Types	Built up	Water areas	Wetland	Forest	Agriculture	GTP	UA	CE
Built up	**78**	0	1	0	1	80	97.50	2.5
Water areas	0	**50**	0	0	0	50	100.00	0
Wetland	0	0	**77**	0	0	77	100.00	0
Forest	0	0	0	**80**	0	80	100.00	0
Agriculture	2	0	2	0	**109**	113	96.46	3.54
Total	80	50	80	80	110	**400**		
PA	97.5	100	96.25	100	99.09	OA=98.5%		
OE	2.5	0	3.75	0	0.91	K=0.981		

Table 3: Contingency matrix, errors, and accuracy assessment using WorldView-2 MS image as source of reference data

LUC Types	Built up	Water areas	Wetland	Forest	Agriculture	GTP	UA	CE
Built up	**77**	0	1	0	1	79	97.47	2.53
Water areas	0	**50**	0	0	0	50	100.00	0
Wetland	0	0	**73**	0	0	73	100.00	0
Forest	0	0	0	**80**	0	80	100.00	0
Agriculture	3	0	6	0	**109**	118	92.37	7.63
Total	80	50	80	80	110	**400**		
PA	96.25	100	91.25	100	99.09	OA=97.25%		
OE	3.75	0	8.75	0	0.91	K=0.96		

Table 4: Contingency matrix, errors, and accuracy assessment using Google Earth image as source of reference data

LUC Types	Built up	Water areas	Wetland	Forest	Agriculture	GTP	UA	CE
Built up	**78**	0	1	0	3	82	95.12	4.88
Water areas	0	**48**	0	0	0	48	100	0
Wetland	0	2	**74**	0	10	86	86.05	13.95
Forest	0	0	0	**79**	0	79	100	0
Agriculture	2	0	5	1	**97**	105	92.38	7.62
Total	80	50	80	80	110	**400**		
PA	97.5	96	92.5	98.75	88.18	OA=94%		
OE	2.5	4	7.5	1.25	11.82	K=0.92		

Table 5: Contingency matrix, errors, and accuracy assessment using GIS layers as source of reference data

LUC types	Built up	Water areas	Wetland	Forest	Agriculture	GTP	UA	CE
Built up	**77**	0	1	0	9	87	88.51	11.49
Water areas	0	**47**	0	0	0	47	100	0
Wetland	0	3	**73**	1	6	83	87.95	12.05
Forest	0	0	1	**76**	0	77	98.7	1.3
Agriculture	3	0	5	3	**95**	106	89.62	10.38
Total	80	50	80	80	110	**400**		
PA	96.25	94	91.25	95	86.36	OA=92%		
OE	3.75	6	8.75	5	13.64	K=0.90		

1.1. *Overall accuracy Kappa coefficient calculations and graphical comparison*

Applying numerically the formula (2) and (3), overall accuracy and Kappa coefficient agreement have been calculated in the table below.

Table 6: Overall accuracy and Kappa coefficient calculation table

Data sources	Resolution	Accuracy measures	Calculations	Results
GeoEye-1	0.5m	Overall accuracy	(78+50+77+80+109)/(78+1+1+50+77+80+2+2+109)=394/400*100	**98.5%**
		Kappa	(400*(78+50+77+80+109))-((80*80)+(50*50)+(77*80)+(80*80)+(113*110))/(400)^2-((80*80)+(50*50)+(77*80)+(80*80)+(113*110))=123710/126110	**0.98**
WorldView-2	2m	Overall accuracy	(77+50+73+80+109)/(77+1+1+50+73+80+3+6+109)=389/400*100	**97.25%**
		Kappa	(400*(77+50+73+80+109)-((79*80)+(50*50)+(73*80)+(80*80)+(118*110))/(400)^2-((79*80)+(50*50)+(73*80)+(80*80)+(118*110))=121560/125960	**0.96**
Google images	3m	Overall accuracy	(78+48+74+79+197)/(78+1+3+48+2+74+10+79+2+5+1+97)=376/400*100	**94%**
		Kappa	(400*(78+48+74+79+97)-((82*80)+(48*50)+(86*80)+(79*80)+(105*110))/(400)^2-((82*80)+(48*50)+(86*80)+(79*80)+(105*110))=116690/126290	**0.92**
GIS layers	Vector model	Overall accuracy	(77+47+73+76+95)/(77+1+9+47+3+73+1+6+1+76+3+5+3+109)=368/400*100	**92%**
		Kappa	(400*(77+47+73+76+95)-((87*80)+(47*50)+(83*80)+(77*80)+(106*110))/(400)^2-((87*80)+(47*50)+(83*80)+(77*80)+(106*110))=113430/126230	**0.90**

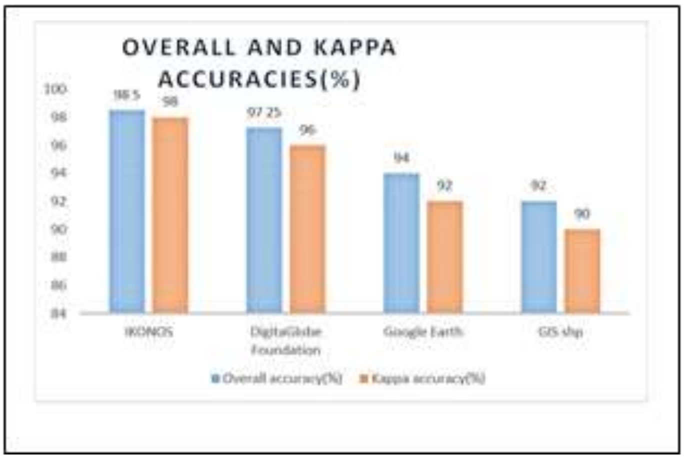

Figure 10: Overall and Kappa accuracies comparison

B. Results discussion

Generally, all reference data sources show excellent accuracy. It is stated that Kappa values of more than 0.80 indicate good classification performance, values between 0.40 and 0.80 indicate moderate classification performance and Kappa values of less than 0.40 indicate poor classification performance[1, 19, 37]. Besides, we do not have a large difference between kappa and overall accuracy. They all range from 90% to 98.5%. This is because the biggest class (Agriculture) accounting for the majority of the map does not show a big difference compared to accurateness of other classes standing alone. Here we can simply take the example of reference data from Google earth image and GIS shape files where correctly classified pixels are estimated under 100 (97/100 pixels for Google Earth images and 95/110 pixels for GIS shape files). Only the difference of 13 pixels and 15 respectively has revealed the difference of 0.2 or 2% between overall and Kappa accuracy. One can straightaway observe 94% (Overall accuracy) and 92% (Kappa accuracy) which makes the difference of 2% for Google Earth images data source, and the effect is the same for GIS layers as reference data source. Let’s glance through GEOEYE-1 and WorldView-2 images. The largest class (Agriculture) didn’t much affect calculated results because only a few pixels from that class were misclassified. Overall accuracy (98.5%) and Kappa Accuracy (98%) are almost equal for reference pixels used from GEOEYE-1 image because only 1 pixel has fallen in built up class and 109 were correctly classified. This is the same case for data from WorldView-2 MS image. Although, wetland also lost 7 pixels out of 80 which contributed to the slightly large range of 97.25% (Overall accuracy) and 96% (Kappa accuracy). Therefore, Overall accuracy is an optimistic index of the classifier performance, even if it is the true agreement of the classification.

The average and percentage summary of overall accuracy (which is basically calculated for the entire study area) may cause unforeseen errors if the classification results are not evaluated wisely. Furthermore, overall accuracy does not reveal if the error is evenly shared by many or all classes and do not indicate if some classes are really negative and some really positive. Therefore, it is required to introduce other forms of errors and accuracy calculations, inclusing: user’s accuracy which corresponds to an error of commission (inclusion error) and Producer's accuracy which corresponds to an error of omission (exclusion error).

Considering results from GIS shape file tests in the agriculture category, which can be applied to other reference data sources results, 86.36% are well labeled or produced accurately (Producer's accuracy), and 13.64 are omitted (Omission errors) in Built up (8.18%) and Wetland (5.15%) from the perspective of the maker of the classified map or reference pixels. Contrarily, in the same category of agriculture 89.62% were correctly classified (User's accuracy) and 10.38% were committed (Commission errors) to the built-up area (2.83%), wetland (4.72%) and forest (2.83%) from the perspective of the operator or user of the classified map. This accuracy means the ratio of pixels correctly identified in agriculture and number appeared to be in the class of agriculture.

V. CONCLUSION AND RECOMMENDATIONS

High-resolution image discernibility of surface features and its application does not ensure the thematic accuracy of classification image since they are totally different yet nonetheless contributes to verify its reliability and in turn quantify the user's confidence through errors and accuracies gauged. Our results showed that four (4) reference data sources examined can be used to assess the accuracy of remotely sensed classified map. Using test data from the field, the results showed that the more the image reveals high spatial resolution, the more it presents high location accuracy. However, it is recommended to go through the following five (5) steps during accuracy assessment performance: (1) Choose reference sources and to make sure it can provide all the information needed, (2) Determine size of reference plots and make sure they match spatial scale for both reference data and classification map, (3) Determine position and number of samples and make sure the landscape is adequately sampled and the sample scheme is verified, (4) Once th data is ready, perform the comparison of reference data with classification results to summarize and quantify statistical estimation, (5) Report the reliability of classification results and deciding for its use. From our results, we can conclude that all the inspected data sources are useful and have a considerable level of classified image accuracy assessment.

Though, as with other data sources, updates and, if possible, field verification, are required using GNSS data that is collected preferably at the same time throughout the day or the week for agriculture study, and throughout the month or the season for the forest and/or desert studies, and urban area studies because the landscape changes with time. It is recommended to focus on soft locations with irregular or fuzzy edges (like rock outcrops and trees and shrub clusters) where it is difficult to distinguish their spectral feature and boundaries using low- or medium-resolution images like those produced by Landsat. The choice of soft edge over hard and and center is optimal and leads to the overall and Kappa accuracy improvement because, the rate of estimating different kinds of objects and features in the center or around regular edges, is too high once the user masters the study area. Although enough reference pixels are required to cover the area and derive any meaningful estimates, through a very few reference pixels are required to deal with time and financial management. In other words, a sizeable enough sample is convenient and must be statistically quantified. In addition, it is recommended to combine reference data with observed data from ancillary data to perform a reliable generalization and class features compactness.

Therefore, El Shayal Smart GIS in conjunction with Google Earth is recommended to gather reference data, as it shows its ability to generate big-rectified high-resolution images [28, 29]. Thus, it is not require physically reaching and checking the area to execute its LUC mapping, since free El-Shayal Smart GIS Editor linked to Google Maps has been a solution to this matter. These images can be downloaded at the scale and spatial resolution of choice. The spatial resolution plays an important role while extracting reference data from images as does projection coordinates system) while extracting testing data from GIS shape files. The higher the image resolution and the more GIS shape files are greatly spatially rectified, the more accurate the level of reference data collected and the less confusion in spectral feature discrimination. This is why GEOEYE-1, WorldView-2 MS, and Google images reveal higher accuracy than GIS shapefiles that were affected by data shift problems in Rwanda's projections.

It is recommended to not add the classified layer in an Arc map table of content while manually collecting reference data. It is not advisable to know which point belongs to which classification type because this can lead to the worst attempt of selecting a class that matches the classification results, which will automatically compromise the accuracy of the assessment performance. only reference data source images and reference points/pixels layers should appear in an Arc map table of content. The classification map is finally opened to perform accuracy assessments through the classified layer and reference pixels combination.

Furthermore, the use of overall and Kappa accuracies should not be considered as redundancy as specified by some researchers, but they should be examined in pairs based on the study type and purpose. Besides, it is better to use and reports different types of accuracy estimations since each of the estimates yields different information istead of only focusing on one estimation that may get an erroneous sense of accuracy. Kappa is generally thought to be a stronger measure than overall accuracy which is considered as a simple percent agreement calculation since Kappa determines the probability that a pixel represents the class for which it has been assigned. For this reason, Kappa accuracy estimates always appear slightly or considerably less than overall estimation, according to the classes errors and accuracy level.

Acknowledgment

The author would like to thank Prof. Zhang X. for his valuable assistance and helpful orientations. The author also wishes to acknowledge the City of Kigali for providing the data used in this research. The research was facilitated by USGS by providing Landsat images at no cost. This research was partially supported by Digital Globe Foundation Grants Program to arrange data grant for this research.

References

[1] K. R. Lillesand TM, Chipman JW Remote Sensing and Image Interpretation, 5th ed., 2004.

[2] G. M. Foody, "Harshness in image classification accuracy assessment," International Journal of Remote Sensing, 29, 3137-3158. , 2008.

[3] S. A. ADEFIOYE, "Quantitative Image Classification Accuracy Assessment Program for Sustainable Geospatial Technology Applications.," 2014.

[4] a. K. J. Mohd Hasmadi Ismail, "Satellite Data Classification Accuracy Assessment Based from reference Dataset," World Academy of Science, Engineering and Technology, International Journal of Environmental, Chemical, Ecological, Geological and Geophysical Engineering, vol. Vol:2, No:3, 2008.

[5] C. R. J., "Sample size for testing differences in proportions for the paired-sample design," Biometrics. 1987 Mar;43(1):207-11., 1987.

[6] G. Gutman, "Assessment of the NASA–USGS Global Land Survey (GLS) datasets," Remote Sensing of Environment 134:249–265 2013.

[7] C. F. B. Gary M. Senseman, and Scott A. Tweddale, "Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping," EN-95/04 1995.

[8] K. G. Russell G. Congalton, Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed., 2008.

[9] J. v. Vliet, "Assessing the Accuracy of Changes in Spatial Explicit Land Use Change Models " presented at the 12th AGILE International Conference on Geographic Information Science 2009 Leibniz Universität Hannover, Germany 2009.

[10] R. G. Congalton, "Accuracy assessment: a critical component of land cover mapping.," 1996.

[11] G. F.-L. Rosenfield, K., " A coefficient of agreement as a measure of thematic classification accuracy.," Photogrammetric Engineering and Remote Sensing, 1986.

[12] G. H. Rosenfield, "The analysis of areal data in thematic mapping experiments.," Photogrammetric Engineering and Remote Sensing, 1982.

[13] G. Foody, "On the Compensation for Chance Agreement in Image Classification Accuracy Assessment," Photogrammetric Engineering & Remote Sensing, vol. Vol. 58. No.10., 1992.

[14] P. S. Thenkabail, " Inter-sensor relationships between IKONOS and Landsat-7 ETM+ NDVI data in three ecoregions of Africa.," International Journal of Remote Sensing 25(2), 389-408. , 2004.

[15] Q. W. Assefa M. Melesse , Prasad S.Thenkabail and Gabriel B. Senay "Remote Sensing Sensors and Applications in Environmental Resources Mapping and Modelling," Sensors 2007, 7, 3209-3241, 2007.

[16] R. L. Demin Xiong, J. Bo Saulsbury, Elizabeth L. Lanzer, Albert Perez, Remote sensing applications for environmental Analysis in transportation planning: : Application to The Washington state i-405 corridor. WA-RD-593-1: Environmental Affairs Office, Washington State Department of Transportation, 2004.

[17] R. G. Congalton, "Thematic and Positional Accuracy Assessment of Digital Remotely Sensed Data," 2005.

[18] S. V. S. a. R. L. Czaplewski, "Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles," REMOTE SENS. ENVIRON., vol. 64:331-344 (1998) 1998.

[19] E. J. O. Charles, "Is 80% accuracy Good Enough?," Michigan Tech Research Institute, Ann Arbor, Michigan 48105, 2008.

[20] E. E. H. JAMES R. ANDERSON, JOHN T. ROACH, and RICHARD E. WITMER "Land Use And Land Cover Classification System For Use With Remote Sensor Data," US Geological Survey Professional Paper No. 964. Washington, DC, 1976.

[21] C. H. Limin Yang, Collin G. Homer, Bruce K. Wylie and Michael J., "An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery " Canadian Journal of Remote Sensing, Coan Raytheon Corporation, U.S. Geological Survey EROS Data Center, Sioux Falls, SD 57198 USA, No date.

[22] G. M. Foody, "Sample size determination for image classification accuracy assessment and comparison," International Journal of Remote Sensing vol. Volume 30, 2009.

[23] G. M. Foody, "Sample Size Determination for Image Classification Accuracy Assessment and Comparison " Shanghai, P. R. China, June 25-27, 2008.

[24] R. Peacock, "Accuracy Assessment Of Supervised And Unsupervised Classification Using Landsat Imagery Of Little Rock, Arkansas," 2014.

[25] A. C. Curtis, "Thematic Accuracy Assessment Procedures, National Park Service Vegetation Inventory,," National Park Service, Natural Resource Program Center atural Resource Report NPS/NRPC/NRR—2010/204, 2010.

[26] G. H. a. F.-L. Rosenfield, K., "A coefficient of agreement as a measure of thematic classification accuracy," Photogrammetric Engineering and Remote Sensing, vol. 52, 1986.

[27] C. G. a. C. Jenkins, "Land cover mapping of Greater Mesoamerica using MODIS data," Can. J. Remote Sensing, vol. Vol. 31, No. 4, 2005.

[28] U. Yusuf. (2013, October 5, 2016). How to Download BIG Rectified Google Earth Imagery to get big High Resolution Image Available: https://www.youtube.com/watch?v=HlpCpi8plXw

[29] M. E. Elshayal, "Elshayal Smart GIS Map Editor and Surface Analysis," Programming of Geographic Information Systems (GIS)-Elshayal Smart GIS Map Editor, 2010.

[30] J. E. Nichol and M. S. Wong, "High resolution remote sensing of densely urbanised regions: a case study of Hong Kong," Sensors (Basel), vol. 9, pp. 4695-708, 2009.

[31] I. Mohd Hasmadi, Alias Mohd,S. and Norizah,K., "Reclassifying forest type to a new forest class based on vegetation and lithology characteristics using geographic information system at southern Johore, Malaysia," INTERNATIONAL JOURNAL of ENERGY and ENVIRONMENT, vol. Volume 2, 2008.

[32] V. Manirakiza, "Promoting inclusive approaches to address urbanisation challenges in Kigali," African Review of Economics and Finance, vol. Vol. 6, No. 1June 2014 pp. 161-180, 2014.

[33] REMA, "Kigali State of Environment and Outlook Report 2013," Rwanda Environment Management Authority, Kigali, Rwanda2013.

[34] P. Avino, D. Brocco, L. Lepore, M. V. Russo, and I. Ventrone, "Remote sensing measurements for evaluation of air quality in an urban area," Ann Chim, vol. 94, pp. 707-14, Sep-Oct 2004.

[35] R. ESRI, How to solve the data shift problem in Rwandan Projections. Kigali, Rwanda, 2012.

[36] W. F. a. B. Kraus, Working with projections and Datum Transformation in ArcGIS: Theory an practical examples. Points Verlag norden Halmstad: 360 gages, richly illustrated paper back 2005.

[37] J. Jensen, Introductory Digital Image Processing, 3rd. ed., 2005.

[38] P. S. Thenkabail, Land Resources Monitoring, Modeling, and Mapping with Remote Sensing, Vol.2. vol. 2, 2016.

[39] J. A. R. Ned Horning, Eleanor J. Sterling, Remote Sensing for Ecology and Conservation: A Handbook of Techniques, 2005.

[40] G. M. Foody, "Status of land cover classification accuracy assessment," Remote Sensing of Environment 80 (2002) 185–201, 2002.

[41] D. G. a. J. Montero, "Determining the accuracy in image supervised classification problems," vol. EUSFLAT-LFA, 2011.

[42] T. A. M. University. (2013, 15 May). Accuracy Assessment of an Image Classification in ArcMap Available: https://www.youtube.com/watch?v=FaZGAUS_Nlo

ISSN: 2409-2770

[43] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed Data," Remote Sensing of Environment, 1991.

[44] B. D. E. a. M. Glass, "The Kappa Statistic: A Second Look," Association for Computational Linguistics, vol. Volume 30, 2004.

[45] M. J. Warrens, "Properties of the quantity disagreement and the allocation disagreement," International Journal of Remote Sensing vol. Volume 36, 2015-Issue 5 2015.

[46] F. Bishop Y.M.M., S.E. and Holland, P.W. , "Agreement as a special case of association. Discrete Multivariate analysis.," Cambridge MA, MIT press, 393 - 400., 1975.

[47] K. Bahadur K.C, "Improving Landsat and IRS Image Classification: Evaluation of Unsupervised and Supervised Classification through Band Ratios and DEM in a Mountainous Landscape in Nepal," Remote Sensing, vol. 1, pp. 1257-1272, 2009.

[48] R. G. B. Congalton, G.S. , "A pilot study evaluating ground reference data collection efforts for use in forest inventory. Photogrammetric Engineering and Remote Sensing.58(12): 1669-1671.," 1992.

[49] R. D. Tortora, "A Note on Sample Size Estimation for Multinomial Populations," The American Statistician, 32, 100-102., 1978.

[50] T. Yamane, "Statistics, An Introductory Analysis," 2nd Ed., New York: Harper and Row, 1967.

[51] J. R. Jensen, " Introductory Digital Image Processing " Prentice-Hall, Englewood Cliffs, NJ, 1986.

[52] W. G. Cochran, "Sampling techniques 3rd ed. New York: Wiley.," 1977.

[53] R. G. G. Congalton, K., "Assessing the accuracy of remotely sensed data: principles and practices," Boca Raton, FL: Lewis Publishers., 1999.

[54] B.-C. a. L. Tai, J. , "Sample size and power calculations for comparing two independent proportions in a ‘negative’ trial," Psychiatry Research, 80, 197-200. , 1998.

[55] G. D. Israel, "Glenn D. Israel2_Determining Sample Size," Fact Sheet PEOD-6, November 1992, Florida cooperative extension services, 1992.

[56] S. V. Stehman, "Comparison of Systematic and Random Sampling for Estimating the Accuracy of Maps Generated From Remotely-Sensed Data," Photogrammetric Engineering & Remote Sensing, vol. vol 58, No. 9, 1992.

[57] J. L. a. L. van Genderen, B.F. , "Testing land use map accuracy," Photogrammetric Engineering and Remote Sensing. 43: 1135-1137., 1977.

Felicien NKOMEJE, born in Rwanda, Eastern Province, Gatsibo District on 17 January 1983. A bachelor degree holder in Geography, majored in Environmental and Planning at University of Rwanda (Former National University of Rwanda) located in Huye, Southern Province, Rwanda-March 2009.

He is currently a post graduate (Master) candidate at Hohai University. He has been a GIS Specialist at Rwanda National Resources Authority (2010-2012), GIS Technician and Cartographer at National Institute of Statictics or Rwanda (2009-2010), and GIS officer at Capital Technology Ltd-Rwanda (2011-2014). His current research interest is Land Use and Land Cover Change modeling and geospatial.