LAND COVER CLASSIFICATION AND CHANGE DETECTION USING LANDSAT IMAGES AND MAXIMUM LIKELIHOOD CLASSIFICATION: THE CASE OF DAVAO CITY, SOUTHERN PHILIPPINES (1996-2016)

Jesselle Lyn C. de Ramos, Joseph E. Acosta

Department of Mathematics, Physics, and Computer Science, College of Science and Mathematics,
University of the Philippines Mindanao, Mintal, Tugbok District, Davao City, Philippines 8000.
Email: jcderamos@up.edu.ph; jeacosta@up.edu.ph

KEY WORDS: Geographic Information System (GIS), image classification, Normalized Difference Built-up Index (NDBI), Normalized Difference Vegetation Index (NDVI), remote sensing

ABSTRACT: Urbanization brings about changes in the land cover of an area from forest or agricultural purposes to commercial, industrial, and residential uses. In the Philippines, land conversion of vegetative areas to built-up areas can be observed in developing cities around the country. With the development of technology, these changes can be observed through images captured by satellites launched to observe the surface of the Earth, with remote sensing and GIS techniques to analyse the data. This study aimed to create NDVI, NDBI and Land Cover maps of Davao City, Southern Philippines to observe changes in the land cover that occurred between 1996 and 2007, 2007 and 2013, and 2013 and 2016. Satellite images of the study area captured by Landsat 5, and Landsat 7 were downloaded from free online platforms and pre-processed to correct scanline errors and subjected to Top of Atmospheric Correction. NDVI and NDBI images were created by using the Red, Near Infrared and Short-wave Infrared bands of each of the selected imagery. Land cover maps of the selected images were generated through image processing methods and Maximum Likelihood Classification following four land cover classes which are Water Body, Bare Land, Built-up Areas, and Vegetation. The accuracy of the classification was assessed by computing the overall accuracy and Kappa Coefficient, with values of 94.29% and 0.92 for the 1996 image, 94.29% and 0.92 for the 2007 image, 89.64% and 0.86 for 2013, and 99.64% and 0.99 for the 2016 image respectively. Results of the change detection indicate an increase in built-up and a decrease in the other land cover classes within 1996 to 2016. In 20 years, water bodies decreased by 47.08%, bare land decreased by 31.55%, built-up areas increased by 33.44%, and vegetation decreased by 2.49%. Field validation may be done in future studies to improve this study. Additional data and analysis in the possible causes of changes in land cover may also prove to be useful in development planning studies.

1. INTRODUCTION

Changes in the Earth's surface can be captured through satellite images with the use of remote sensing technology (Shalaby and Tateishi, 2007). Using satellites that capture multi-spectral images of the surface of the planet, remote sensing acquires information and extract spectral, spatial, and temporal features of an object, area, or phenomenon, including land cover classification without physical contact with them (Gandhi et al., 2015).

In the Philippines most of the changes in land cover was caused by urbanization, and because of it, land areas are being converted to residential and commercial areas from their original state which for most is mainly forest cover. In a study by Janiola and Puno (2018), remote sensing proved to be a useful tool in monitoring land use and land cover. In the Allah valley landscape located in Southern Mindanao, conversion of forest areas to commercial and industrial use caused a constant decrease in the forest cover of the area (Janiola and Puno, 2018). Estoque and Murayama (2011) linked urbanization as a primary factor in the conversion of other land classes to built-up areas in Baguio City, with a 1909.26 ha increase in built-up in 21 years.

Based on the study by Soriano et al. (2019), urban sprawl caused a significant change in the land cover of Mount Makiling Forest Reserve watersheds between 1992, 2002, and 2015. By monitoring land cover change, the researchers observed a 117% increase in built-up areas from 1992 to 2015, with 3.4% of the buffer zone classified as built-up area in 2015. From their findings, the researchers suggested that regular monitoring of ecologically valuable areas using remote sensing and GIS techniques could help in implementing policies that aim to protect and conserve these areas.

As presented in the Comprehensive Land Use Plan of Davao City for 2013-2022, a significant transformation in land use occurred between 1994 and 2011 (City Government of Davao, 2012). A high demand for land area for residential use resulted from the increase in population. This resulted to a 103% increase in residential areas in 2011 compared to 1994 and by 206% for commercial use. With the use of remote sensing and GIS, these changes can be analyzed to aid in the development and conservation plans of the city.

This study aimed to detect changes in the land cover of Davao City for 1996 to 2016 using land cover maps derived from satellite imagery. Specifically, the study is aimed to: (i) generate Normalized Difference Vegetation Index (NDVI) and Normalized Difference Built-up Index (NDBI) images; (ii) create Land Cover Classification maps; and (iii) detect changes in the land cover of the study area using the generated land cover, NDVI, and NDBI maps.

2. MATERIALS AND METHODS

2.1 Study Area

The area of interest for the study was Davao City, Philippines (7.1907° N, 125.4553° E). The city has an area of 244,000 hectares, which is approximately 8% of Region XI (City Government of Davao, 2012).The city government lists three major groups of the area's land use capability and environmental classification: agriculture, conservation areas, and resource conservation.

2.2 Software Used

Conversion of DNs to Top of Atmosphere (TOA) reflectance was done in QGIS Desktop 3.2.0. Scan line error correction, image classification, and generating error matrices for classification accuracy assessment were all done using ArcMap 10.2.1 in ArcGIS licensed under the Sago 2.1 Perpetual License.

2.3 Data

Images on the study area collected by Landsat 5 and Landsat 7 satellites were downloaded from the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center and National Aeronautics and Space Agency (NASA) Earthdata which are both platforms that provide free and publicly accessible satellite images. The downloaded images were evaluated for usability based on their cloud cover percentage. Images captured on 1996, 2007, 2013, and 2016 were selected for the implementation of the methodology considering their minimal cloud cover. The selection of images that have different intervals between them means that change occurred between 11 years, 6 years, and 3 years may be observed during change detection.

A map showing the boundary of the study area was obtained from the GeoSAFER Project – UP Mindanao. The Comprehensive Land Use Plan of Davao City which was used as a reference for the analysis was downloaded from the city government's website prior to the conduct of this study.

2.4 Pre-processing of Satellite Images

Scan line error correction was done on affected Landsat 7 images to improve the usability of the downloaded data. This was done using the Landsat tool custom developed and used by Daniels (2010) for ArcMap. After scan line correction, Top-of-Atmospheric (TOA) correction was done on the images to filter at-sensor radiance and eliminate additional radiation obtained from the atmosphere. TOA correction was done using the pre-processing option in the Semi-Automatic Classification Plugin in QGIS. All bands of the selected images were loaded for correction, using information contained in their respective metadata.

2.5 Vegetation Index Mapping

Land cover change was also observed according to the changes in vegetative cover. For this, Normalized Difference Vegetation Index (NDVI) was applied on the obtained satellite images. NDVI monitors the condition of vegetation and vegetation health in an area utilizing the uniqueness in shape of the reflectance curve of vegetation (Zha et al., 2003). NDVI is calculated as the difference between the near-infrared reflectance value and the red reflectance value of an image normalized with their sum (Nageswara Rao et al., 2005). Hence, the following formula is derived (Zha et al., 2003):

$$NDVI = \frac{(NIR-R)}{(NIR+R)} \qquad [1]$$

where,
NIR = Near Infrared (band 4 for both Landsat 5 TM, and Landsat 7 ETM+)
R = Red (band 3 for both Landsat 5 TM, and Landsat 7 ETM+).

The NDVI was implemented by using Map Algebra under the Spatial Analyst tools in ArcMap 10.2.1 following the formula stated above.

2.6 Built-up Index Mapping

The Normalized Difference Built-up Index (NDBI) was also considered as a way of determining land cover change. The following formula will be followed in computing the NDBI (As-syakur et al., 2012):

$$NDBI = \frac{SWIR-NIR}{SWIR+NIR} \qquad [2]$$

where,
SWIR = Short-wave Infrared (band 5 for both Landsat 5 TM, and Landsat 7 ETM+)
NIR = Near Infrared (band 4 for both Landsat 5 TM, and Landsat 7 ETM+).

The NDBI was implemented by using Map Algebra under the Spatial Analyst tools in ArcMap 10.2.1 following the formula stated above.

2.7 Land Cover Classification

Land cover classes must be identified to analyze changes in land cover. This study roughly followed the definition of land cover classes in remote sensing studies set by Anderson et al. (1976) and used four major land cover classes: Water Body, Bare Land, Built-up Areas, and Vegetation. Having done so ensured that classes were well-separated in order to avoid confusion.

Training samples were selected per class on each image to serve as the basis in assigning each pixel to a specific class. Polygons were created per land cover class to serve as the training sample. A signature file was created using the polygons and was then used for classification.

2.8 Maximum Likelihood Classification

In remote sensing data, maximum likelihood classification works by classifying each cell based on the highest probability of being a member and spectral pattern represented in class mean vectors and covariance matrices (Soriano et al., 2019). The main assumption for the Maximum Likelihood Classification algorithm is that the probabilities for all classes are equal and the training samples in each band are normally distributed. As shown in ERDAS (1997), the formula in solving for the weighted distance (likelihood) D is as follows:

$$D = \ln(ac) - [0.5\ln(|Covc|)] - [0.5(X - Mc)T(Covc - 1)(X - Mc)] \qquad [3]$$

where
D = weighted distance (likelihood)
C = a particular class
X = the measurement vector of the candidate pixel
Mc = the mean vector of the sample of class *c*
Ac = percent probability that any candidate pixel is a member of class *c*
(defaults to 1.0, or is entered from *a priori* knowledge)
Covc = the covariance matrix of the pixels in the sample of class *c*
|*Covc*| = determinant of *Covc* (matrix algebra)
Covc -1 = inverse of *Covc* (matrix algebra)
Ln = natural logarithm function
T = transposition function (matrix algebra)
Where the pixel for evaluation is assigned to the class c that generates the lowest D (ERDAS Inc, 1997).

The Image Classification tool of ArcMap 10.2.1 was used in the classification, following the Maximum Likelihood method and using the signature file generated in the previous step.

2.9 Accuracy Assessment

The resulting land cover maps were subjected to accuracy assessment using Overall Accuracy, Cohen's Kappa, and other measures of image classification accuracy through generating an error matrix in ArcMap 10.2.1.

Overall accuracy is computed as the total correct or the sum of the major diagonal of the matrix divided by the total number of pixel in the matrix (Soriano et al., 2019).
The Kappa Coefficient is computed by (Jensen, 1996):

$$K = N\sum(X_{ij}) - \sum(row_i\ total)(col_j\ total)/N^2 - \sum(row_i\ total)(col_j\ total) \quad [4]$$

where:
N – total number of observations
X_{ij} – sum of the major diagonal
row_i – marginal total for row_i
col_j – marginal total for $column_j$

The K interpretation values range from 0 to 1, the closest the value is to 1 the more acceptable the classification (Soriano et al., 2019).

All accuracy tests were done in Microsoft Excel using the confusion matrix generated in ArcMap. Classified images with Overall Accuracy values greater than 80% and Kappa Coefficient value of 0.80 or higher were accepted for change detection, while those with lower values were set to be classified again using a new training sample and signature.

The lack of ground truth data for images taken on past years (1996, 2007, 2013, and 2016) meant that this study will solely rely on the accuracy of the classification process. Ground accuracy was assumed based on the classification accuracy.

2.10 Change Detection

Image differencing was used in identifying the differences between the image composites of: 1996 and 2007, 2007 and 2013, 2013 and 2016, and 2016 and 2018. This involves subtraction of the older image from the newer image composite. The method was followed in taking the differences in the NDVI and NDVI between the specified time periods. This involves subtracting the older NDVI/NDBI image from the newer NDVI/NDBI image. For this purpose, the Raster Calculator under the Spatial Analyst tool in ArcMap 10.2.1 was used. Post classification comparison was done to detect changes in the generated land cover maps.

3. RESULTS AND DISCUSSION

3.1 Preprocessing of Data

The downloaded images were subjected into the necessary pre-processing methods. Landsat 7 images with scanline errors were corrected and all the bands of all images collected were cleaned of unnecessary reflectance using the TOA Correction. Images to be used further in the study were selected through an evaluation of their cloud cover percentage. Then, NDVI, NDBI, and land cover maps were generated.

3.2 Normalized Difference Vegetation Index (NDVI) and Normalized Difference Built-up Index (NDBI)

The NDVI and NDBI have values ranging from -1 to 1, values close to -1 means a low concentration of vegetation while values close to 1 means a high concentration of vegetation. For easier interpretation of the indices, color gradients were used to demonstrate high and low values: for the NDVI, low values in hues of red while high values are in hues of green, while for the NDBI low values from -1 is shown in hues of green while high values until 1 in hues of red.

Note that in most of the images, cloud cover obstructs the actual vegetation of the study area, and hence may be interpreted in this study as inconclusive. For example, in the 1996 image, cloud cover was classified as low vegetation in a significant area of the image.

Discarding areas with cloud cover, other areas with low vegetation can be observed in areas where water is present (rivers, streams, etc.) and in areas with dense urban activity (commercial and residential areas, etc.) in the 1996 image as shown in Figure 1. High vegetation values can be found in areas with dense vegetation (forests, grasslands, etc.) and with visibly less urban presence.

Figure 1 also shows the NDVI image for 2007. Similar to the 1996 NDVI image, low values can also be observed in areas with high urban activity. Compared to the previous image, there is notably more areas that lean to neutral areas and low vegetation value areas.

More low value to neutral areas can be seen in the 2013 NDVI image also in Figure 1. The area with cloud cover which can be seen on the 2007 image was classified as areas having high to neutral vegetation in the

2013 NDVI image where that area was already free of cloud cover. The trend continues for the 2016 NDVI image still in Figure 1.

The observed trend in the NDVI images is also reflected in the NDBI, as shown on Figure 2. Urban development can be seen growing in each year, as observed in the increase of areas with neutral to high built-up values through the years.

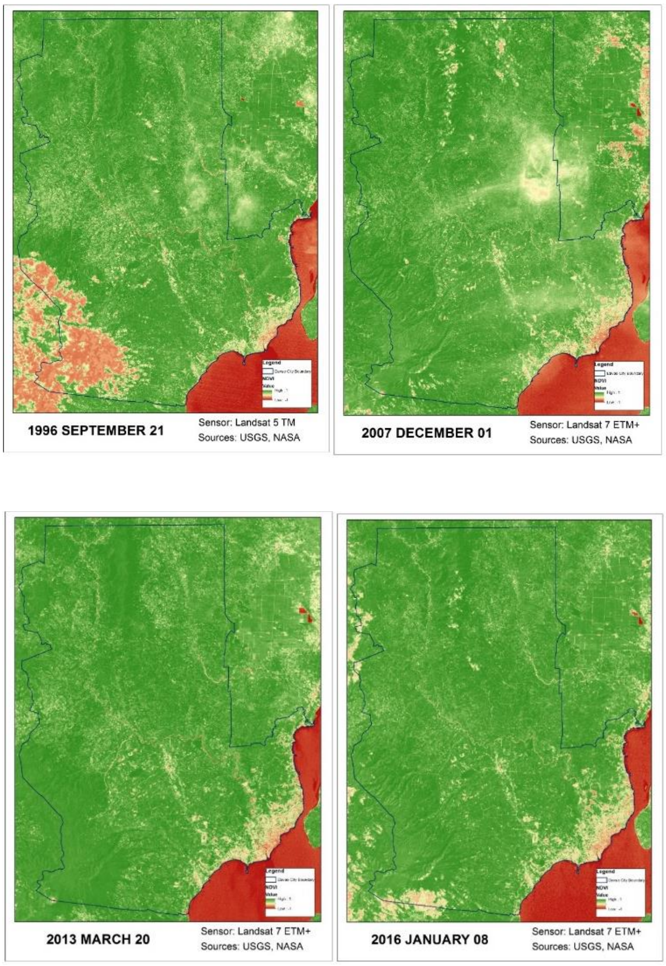

Fig. 1 (Left to right, top to bottom) NDVI maps generated using bands 3 and 4 of Landsat 5 TM and Landsat 7 ETM+ images captured on September 21, 1996, December 01, 2007, March 20, 2013, and January 08, 2016.

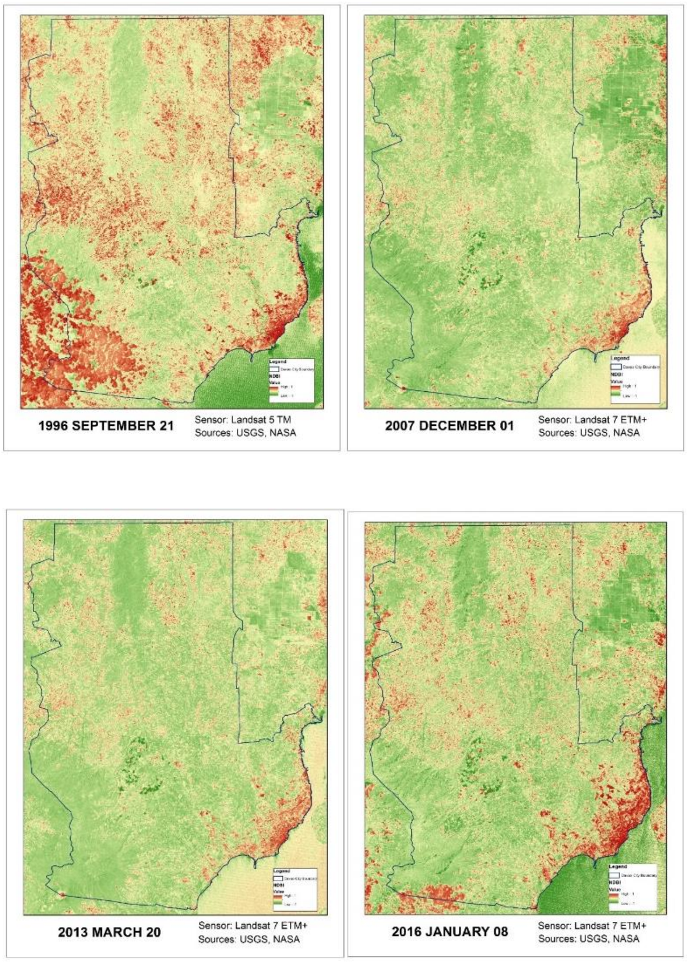

Fig. 2 (Left to right, top to bottom) NDBI maps generated using bands 3 and 4 of Landsat 5 TM and Landsat 7 ETM+ images captured on September 21, 1996, December 01, 2007, March 20, 2013, and January 08, 2016.

3.3 Land Cover Classification and Change Detection

Classified pixels may contain more than one land cover class in actual ground data but may only be processed as the dominant land cover class due to the constraint of the image's spatial resolution. It is also important to note that the difference in cloud cover of the images plays a huge factor in the classification. Since cloud cover was not considered as a feature class in the classification process, the areas on the images which were covered with clouds were classified into a class which the algorithm identified as having a similar signature. Inconsistencies in the signature of clouds and cloud shadows may be caused by the variations in type and volume of clouds present in an image.

For the 1996 image, a major part of built-up and water body is actually cloud cover, as can be seen in Figure 3, which has caused a significant addition to the actual area of each mentioned class.

Figure 3 also shows the land cover map for 2007 and it can be seen that cloud cover still affects the classification. In this case, cloud cover was classified as water bodies (middle and lower right area), and that caused a significant increase in area compared to the 1996 image.

The 2013 image does not have any visible cloud cover that may significantly affect classification, as shown in Figure 3. The cloud cover classified as water body in 2007 was classified mostly as vegetation in 2013 where that area was already free of cloud cover.

The 2016 land cover map in Figure 3 has little cloud cover on its upper and bottom left and upper middle part however the volume is not enough to have a significant effect on the area of the class it is classified into, which is bare land.

Table 1 shows the area of each land cover class for each of the images studied.

Table 1. Area (in thousand hectares) of each land cover class per studied year.

Land Cover Class	Year			
	1996	2007	2013	2016
Water	3.7555	37.2974	12.7563	1.9876
Bare Land	7.2930	3.1853	23.2261	4.9922
Builtup	26.8158	10.9620	12.8874	35.7835
Vegetation	196.5055	182.9250	185.5000	191.6065
TOTAL	234.3697	234.3697	234.3697	234.3697

Table 2 shows the differences of the area of each class between the years studied. Note that the significant increase in water body between 1996 and 2007 is due to the cloud cover on 2007 being classified as water body.

Table 2. Differences of each land cover class between two images studied (in thousand hectares).

Land Cover Class	2007-1996	2013-2007	2016-2013
Water Body	33.5419	-24.5411	-10.7687
Bare land	-4.1077	20.0408	-18.2340
Built-up	-15.8538	1.9254	22.8961
Vegetation	-13.5804	2.5749	6.1066

Meanwhile, the significant decrease in built-up between the same years is due to the cloud cover on 1996 being classified as built-up. The increase in bare land in 2013 may be attributed to the classification of unused planting area which can be classified as vegetation during planting season and bare land or built-up during harvest season.

3.4 Accuracy Assessment

The classified images were subjected into accuracy assessment to check the efficiency of the classification. Overall accuracy and Kappa coefficient were selected as the primary measure of accuracy for the classification process.

The classified 1996 composite image had an overall accuracy of 94.29% and a Kappa coefficient of 0.92. The 2007 image had similar results with an overall accuracy of 94.29% and a Kappa Coefficient of 0.92. The overall accuracy and Kappa coefficient of the classified 2013 image were 89.64% and 0.86 respectively. Finally, the 2016 image had an overall accuracy of 99.64% and a 0.99 Kappa coefficient. The overall accuracy and Kappa coefficient of all images were greater than the set standard, which was 80% and 0.8 respectively. Thus, the classification for all time periods was accepted and were then subjected to change detection.

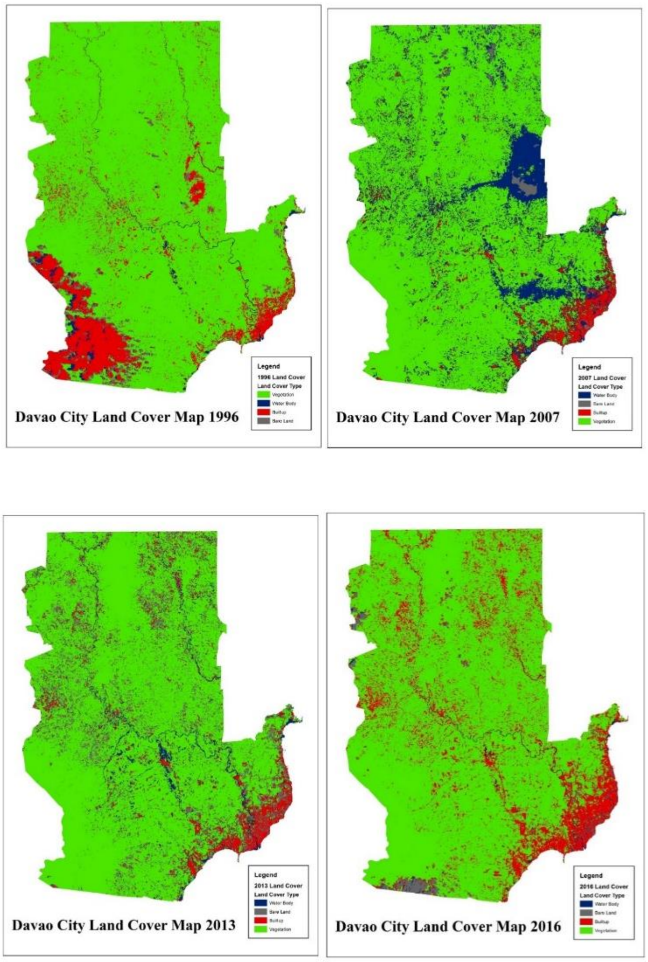

Fig. 3 (Left to right, top to bottom) Resulting Land Cover Maps of Davao City on 1996, 2007, 2013, and 2016 from classification of image composites made from Landsat 5 TM and Landsat 7 ETM+ images captured on the years mentioned earlier.

3.5 Change Detection

Change detection on the NDVI and NDBI images was done using image difference. Change detection for image differencing as well as in post classification comparison in land cover maps followed a specified time period comparison: 1996 vs 2007, 2007 vs 2013, and 2013 vs 2016.

The NDVI difference values are still in an index scale, where high values mean an increase in vegetation and low values indicate a decrease. A color gradient was used to highlight the difference between high and low values, as well as areas that showed no change: low values are in red, neutral or no change values are in yellow, and high values are in blue.

In a span of 11 years, Figure 4 shows the changes in vegetation that occurred between 1996 and 2007. Disregarding the cloud cover, high values which indicate an increase in vegetation can be seen in the northern areas. Areas in the south show a visible decrease in vegetation, as indicated by their low values on the scale, from this we may hypothesize that this decrease was caused by the development of commercial and residential areas in the downtown part of the city.

Only a small part of the 2007-2013 image difference shows areas with decreasing vegetation, as also shown in Figure 4. Disregarding the cloud cover, there are also only a few areas which show an increase in vegetation. Majority of the area are in light hues of blue, indicating a close approach to neutral or no change. This may mean that no significant changes occurred in a span of 6 years, at least not as much as the change shown in the previous image which spanned 11 years.

Within a 3-year interval, only a few changes can be observed in the 2013-2016 image difference as seen on

Figure 4. Again, disregarding the cloud cover, there are only a few spots that show a decrease in vegetation which all can be found on the southern areas. Areas with increased vegetation are fewer than those with decreased, and can be observed in the lower middle area of the city.

Similar with the NDVI, the NDBI difference values are also in an index scale, where high values mean an increase in built-up and low values indicate a decrease. A color gradient was used to highlight the difference between high and low values, as well as areas that showed no change: low values are in red, neutral or no change values are in yellow, and high values are in blue.

Disregarding the cloud cover, within an interval of 11 years, a significant increase in built-up can be observed all over the city as seen in the difference image between 1996 and 2007 in Figure 5. This increase may be caused by infrastructure development in the form of commercial establishments and residential units such as subdivisions and condominiums. Meanwhile, there are also notable decrease in built-up in the northern areas, which may be explained by different scenarios such as the difference in the time that the images were taken were probably two end points of the planting season. Before planting season, land may be identified as built-up or bare land, while during planting season it may be identified as vegetation.

Within a six year interval in 2007 to 2013, increase in built-up was still observed especially in the downtown areas. Few places exhibited a clear decrease in built-up while most were into not significantly low values and are in hues of light red to the neutral no change yellow. This can be seen in the 2013 – 2007 NDBI difference image also in Figure 5.

The steady increase in built-up can be observed in the 2013 – 2016 image still in Figure 5. Areas in the northern part of the city show increasing built-up as well, not just in the southern part where the center for economic activity can be found. Decreasing built-up can still be observed in the northeast and southwest areas of the city. But majority of the image lies in the neutral values, indicating no change.

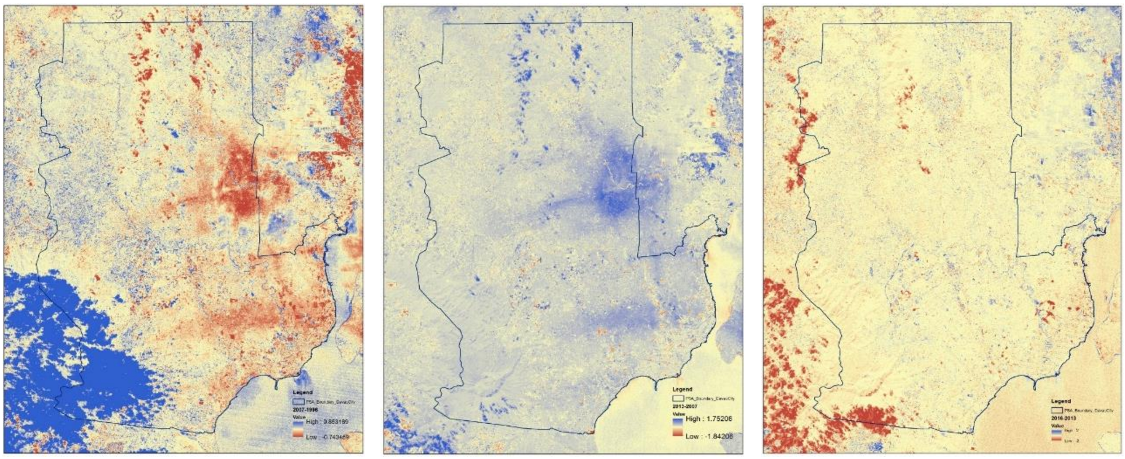

Fig. 4 (Left to right) Computed NDVI Difference for 1996 and 2007, 2007 and 2013, and 2013 and 2016.

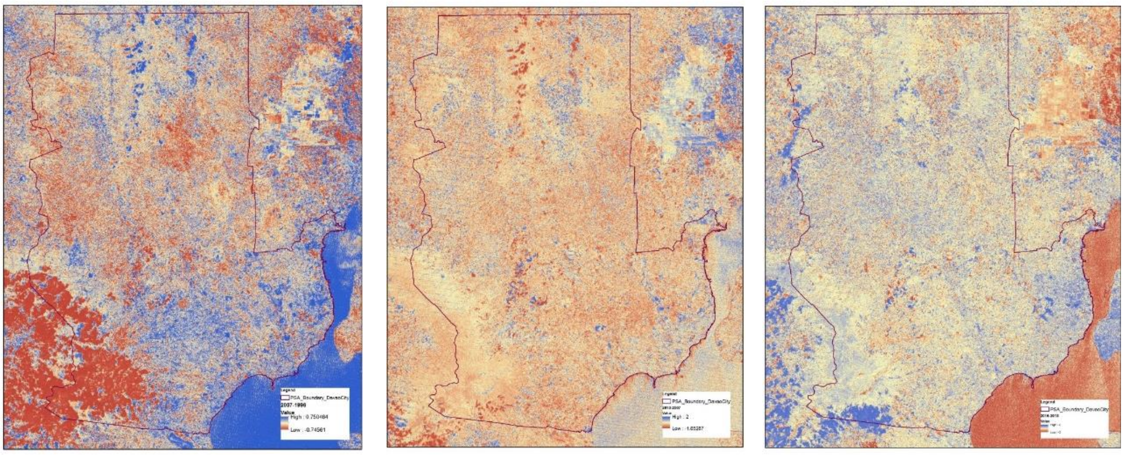

Fig. 5 (Left to right) Computed NDBI Difference for 1996 and 2007, 2007 and 2013, and 2013 and 2016.

4. CONCLUSIONS

Change detection through image differencing and post classification comparison of the NDVI, NDBI, and derived land cover maps of the study area showed that different intervals between time periods had different levels of change on the landscape of the area. From 1996 to 2016, water bodies decreased by 47.08%, bare land decreased by 31.55%, built-up areas increased by 33.44%, and vegetation decreased by 2.49%. The resulting differences between images of different years were observably influenced by presence of cloud cover, but ultimately, results indicate an increase in built-up areas as time progressed.

For future studies, land cover classes can be more specified to extract more specific data. Classes can be increased, to differentiate land use or a specific land cover such as grasslands, dense forests, croplands, etc. Other factors such as socioeconomic and/or anthropogenic activities may be able to provide the needed evidences to point out the causes of changes that can be observed from studies such as this. Incorporating other data may prove to be useful in creating development and sustainability plans in the future, as well as the application of the methods in different fields such as agriculture and disaster risk reduction and management.

5. REFERENCES

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., 1976. A land use and land cover classification system for use with remote sensor data. Geol. Surv. Prof. Pap. 964 34.

As-syakur, A.R., Adnyana, I.W.S., Arthana, I.W., Nuarsa, I.W., 2012. Enhanced built-up and bareness index (EBBI) for mapping built-up and bare land in an urban area. Remote Sens. 4, 2957–2970.

City Government of Davao, 2012. Comprehensive Land Use Plan (2013-2022).
Daniels, R.C., 2010. Using ArcMap to Extract Shorelines from Landsat TM & ETM+ Data 1–23.

ERDAS Inc, 1997. ERDAS Field Guide. Imagine 645.
Estoque, R.C., Murayama, Y., 2011. Spatio-temporal urban land use / cover change analysis in a hill station : the case of Baguio City, Philippines. Procedia - Soc. Behav. Sci. 21, 326–335.

Gandhi, G.M., Parthiban, S., Thummalu, N., Christy, A., 2015. Ndvi: vegetation change detection using remote sensing and GIS - a case study of Vellore District. In: Procedia Computer Science. Elsevier Masson SAS, pp. 1199–1210.

Janiola, M.D.C., Puno, G.R., 2018. Land use and land cover (LULC) change detection using multi- temporal landsat imagery : A case study in Allah Valley Landscape in Southern , Philippines. J. Biodivers. Environ. Sci. 12, 98–108.

Jensen, J.R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective. Prentice Hall Inc., Upper Saddle River, NJ, USA.

Nageswara Rao, P.P., Shobha, S.V., Ramesh, K.S., Somashekhar, R.K., 2005. Satellite-based assessment of agricultural drought in Karnataka State. J. Indian Soc. Remote Sens. 33.

National Aeronautics and Space Agency, n.d. NASA Earthdata [WWW Document]. URL https://earthdata.nasa.gov (accessed 2.4.19).

Shalaby, A., Tateishi, R., 2007. Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt. Appl. Geogr. 27, 28–41.

Soriano, M., Hilvano, N., Garcia, R., Hao, A.J., Alegre, A., Tiburan, C., 2019. Land use / land cover change detection and urban sprawl analysis in the Mount Makiling forest reserve watersheds and buffer zone, Philippines. Environments 6, 1–20.

United States Geological Survey (USGS), n.d. USGS EROS [WWW Document]. URL https://www.usgs.gov (accessed 2.4.19).

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int. J. Remote Sens. 24, 583–594.