

LAND COVER MONITORING OF LAGUNA LAKE WATERSHED USING MODIS NDVI DATA

J.M. Medina^{1,*}, A.C. Blanco^{1,2}, C.G. Candido¹

¹Training Center for Applied Geodesy and Photogrammetry, University of the Philippines Diliman, 1101 Quezon City, Philippines - jmmedina@up.edu.ph, chriscandido93@gmail.com

² Department of Geodetic Engineering, University of the Philippines Diliman, 1101 Quezon City, Philippines - acblanco@up.edu.ph

KEY WORDS: LULC, vegetation index, ISODATA, spacetime cube, hot spot analysis, MAD, canonical correlation

ABSTRACT:

Land use and land cover monitoring is an important component in the management of Laguna Lake watershed due to its impacts on the lake's water quality. Due to limitations caused by cloud cover, satellite systems with limited revisit capability fail to provide sufficient data to more effectively monitor the land surface. Normalized difference vegetation index (NDVI) derived from MODIS image data were used to generate land cover maps for the years 2001, 2005, 2009, 2013, and 2017. These were produced by classifying ISODATA classes using annual NDVI profiles, which resulted in land cover classes, namely, agricultural land, built-up, forest, rangeland, water, and wetland. The resulting maps were post-processed using multi-variate alteration detection (MAD), resulting in multi-temporal land cover maps with improved overall accuracies and kappa coefficients that indicate moderate agreement with ground truth data. Spatiotemporal hot spot analysis was also performed using NDVI data from 2001 to 2017 to identify vegetation hot spot areas, where clustering of low NDVI values were observed over the years. Results showed an increasing trend in built-up areas accompanied by decreasing trends in water and wetland areas, indicating impacts caused by land reclamation and expansion of residential subdivisions near the lakeshore. The decrease in total vegetation area from 2001 to 2017 could be attributed to conversion of land to built-up surface. Vegetated areas in identified hot spots decreased from 41% in 2001 to 19% in 2017. This suggests that vegetation cover in these hot spots was converted to non-vegetated surface during the time period studied.

1. INTRODUCTION

Remote sensing satellites acquire tremendous amount of data which have been useful in monitoring the Earth's surface. Time-series data coming from high temporal resolution sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) have been used in numerous studies to monitor and assess land cover using vegetation datasets such as the normalized difference vegetation index (NDVI). The common finding is that using MODIS time-series data for land cover studies is cost- and time-efficient for mapping large areas (Nyamekye, et al., 2014).

Laguna Lake is the largest inland body of water in the Philippines. It is situated in the heart of the CALABARZON-Metro Manila region, which consists of six provinces and is one of the most densely populated areas in the country. These geographical and demographic characteristics make land cover monitoring a vital component in resource management of the lake as land use and land cover have an impact on the lake's overall quality, among other concerns.

Nationwide land cover maps, which are commonly used in different agencies, are released every five years in the Philippines. The long time between successive releases makes monitoring of land cover changes difficult especially for purposes of resource management. This concern is important when dealing with highly dynamic and fast-changing systems such as the Laguna Lake watershed. Because of the large extent of land area covered and the many different human activities happening in it, there is a need to consider other datasets which can provide information faster at an acceptable level of accuracy. Time-series MODIS NDVI offer an alternative to address this issue.

2. MATERIALS AND METHODS

2.1 Study Area

Laguna Lake has a surface area of about 900 sq km, making it the largest inland waterbody in the Philippines and third largest in Southeast Asia (Laguna Lake Development Authority, 2017). Otherwise known as Laguna de Bay, the lake has an average depth of about 2.5 m (Santos-Borja, Nepomuceno, 2003). It is used for various purposes such as aquaculture, navigation, flood control, and source of water for irrigation, industrial cooling, and drinking.

With a shoreline of about 285 km, the lake is divided into four regions namely West, Central, South, and East Bays, each with unique water quality characteristics as a result of effluents coming from more than 100 tributaries draining into the lake from the surrounding area. Its only outlet is the Pasig River which drains into Manila Bay. Because of such hydrologic set-up, backflow occurs during the dry season when lake water level is lower than the mean sea level. This results in the intrusion of saline water which enters the West Bay through the Pasig River making the lake water brackish (Santos-Borja, Nepomuceno, 2003).

The watershed area of Laguna Lake is about 2,980 sq km and geographically covers, in part or in full, 59 towns in Metro Manila and the Provinces of Laguna, Rizal, Cavite, Batangas, and Quezon (Santos-Borja, Nepomuceno, 2003). The urban towns in Laguna and Metro Manila are situated in the west portion of the watershed while the north and east are characterized by a mixture of built-up and vegetation cover. Agricultural lands are mostly found in the east and south areas. For management purposes, Laguna Lake watershed is divided

into 24 sub-watersheds. Figure 1 shows a map of the watershed area.

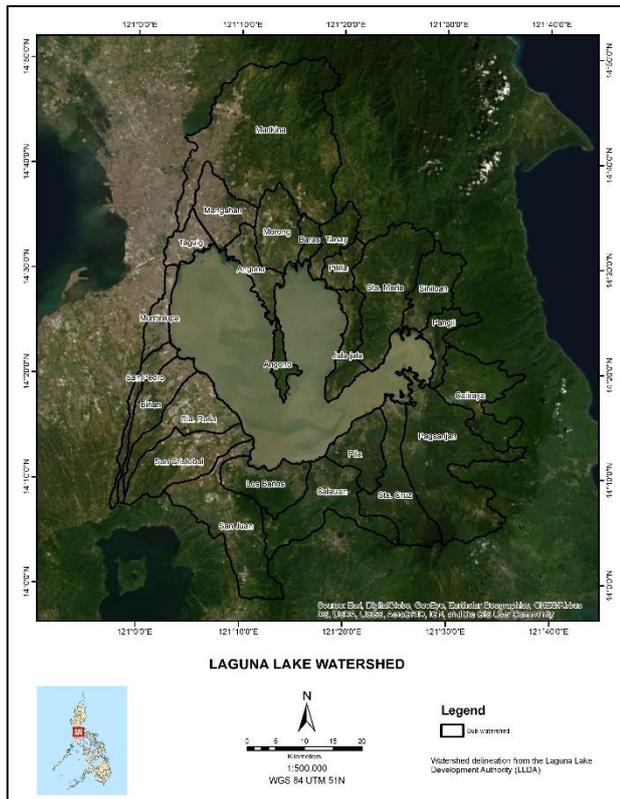


Figure 1. Laguna Lake watershed area

2.2 MODIS NDVI Data

Moderate Resolution Imaging Spectroradiometer (MODIS) is a remote sensing instrument onboard the Terra and Aqua satellites. These two satellites acquire complete coverage of Earth observation data every one to two days in sun-synchronous near-polar orbits, with Terra making an overpass from north to south across the equator in the morning and Aqua making an overpass in reverse direction in the afternoon (MODIS Moderate Resolution Imaging Spectroradiometer, n.d.)

The MODIS sensor captures data in 36 spectral bands with wavelength ranging from 0.4 to 14.4 μm available in three spatial resolutions. Bands 1 (red; 620-670 nm) and 2 (near-infrared; 841-876 nm) are at 250-m resolution. Bands 3 to 7 (459-479 nm, 545-565 nm, 1230-1250 nm, 1628-1652 nm, 2105-2155 nm, respectively) are at 500-m resolution. Finally, Bands 8 to 36 have 1-km resolution (MODIS Moderate Resolution Imaging Spectroradiometer, n.d.). Bands 1 to 7 are commonly referred to as land bands as these are used for remote sensing applications of the land surface (Carroll, et al., 2010).

The normalized difference vegetation index (NDVI) is a remote sensing product used to measure vegetation health and density. It has a value ranging from -1 to 1, where lower values indicate sparse and/or less healthy vegetation while higher values denote denser and/or healthier vegetation. It is based on measured reflectance from red and near-infrared bands of a satellite image and is defined by the following equation:

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

MOD13Q1 Version 6 data are a product derived from image data acquired by Terra MODIS sensor. The dataset is a Level 3 MODIS product generated as 16-day composites at 250-m spatial resolution. To produce the composite data, the algorithm chooses the best pixel from all acquisitions within the 16-day period based on the following criteria: (i) minimum cloud contamination; (ii) low view angle; and (iii) highest NDVI value. The dataset is in sinusoidal projection and is available from February 2000 up to present (Didan, 2015). Accuracy of MODIS NDVI data is within ± 0.025 . This value represents the accurate retrieval of top-of-canopy and nadir VI values for high quality observations (MODIS Land Team Validation, n.d.).

2.3 Pre-processing

MOD13Q1 data covering Luzon island from 2001 to 2017 were downloaded from NASA's Earthdata site. Each year consists of 23 granules (or scenes), with each granule representing 16-day composite data consisting of 12 layers including VI layers, surface reflectance in MODIS Bands 1, 2, 3, and 7, and pixel quality layers, among others.

The NDVI layer was extracted from each of the downloaded MODIS data and re-projected to WGS 84 UTM Zone 51N. The NDVI layers acquired during the same year were stacked in order from Julian date-of-year 001 (January 01-16) to 353 (December 19-31) to produce the NDVI dataset for each year. Subsets of the datasets were generated following the extent of Laguna Lake watershed with 500-m buffer.

As the data are already a Level 3 product, these have already undergone pre-processing to minimize effects of the atmospheric elements, such as clouds, in the data. However, to further minimize cloud contamination effects, Savitzky-Golay (SavGol) filtering was done on each of the NDVI stacks. The SavGol filter is based on the study of Savitzky and Golay (1964) which presented an approach to smoothen a set of consecutive values, or a spectrum, by means of a simplified convolution using least squares fitting within a moving window. Weights are given such that the observations within the filter window are fitted to a polynomial whose degree is equal to the weight (Chen, et al., 2004).

Half-width smoothing window m and degree of polynomial d are required parameters for SavGol filter. For this study, the value of (m, d) used was $(4, 6)$, as it is the optimal pair of parameters determined by Chen, et al. (2014) to provide the best fit. For each pixel location in the stacked NDVI layers, the values were taken by SavGol filter and fitted to generate the smoothened dataset for that location. The results of this filtering are smoothened NDVI datasets for Laguna Lake watershed from 2001 to 2017.

2.4 Land Cover Classification

Smoothened NDVI datasets were used to generate land cover classes using Iterative Self-Organizing Data Analysis Techniques (ISODATA) algorithm. ISODATA is an unsupervised classification technique in which pixels are clustered based on mean distances between pixels and class centroids and standard deviations. Splitting of clusters is done when the maximum standard deviation exceeds a user-specified value and when the number of pixels in the cluster is twice a specified minimum number. Clusters are merged if the distances between their centroids are less than the minimum distance specified (Ball, Hall, 1965). New class distances and standard

deviation for each cluster are calculated after each iteration and new clustering is performed. The iteration stops when the maximum number of iterations is reached, the maximum number of ISODATA classes specified is achieved, and/or no classes can be further split or merged (Dhodhi, et al., 1999).

The threshold values for class standard deviation and class distance were determined by calculating these statistics from annual average NDVI. Land cover and land use data for the study area released by the National Mapping and Resource Information Authority (NAMRIA), the Philippines' national mapping agency, as of 2010 were used as arbitrary classes for the computation of zonal statistics. The mean values and standard deviations for the zones were compared and used to derive the minimum class difference (distance) and maximum standard deviation used as input parameters in ISODATA classification. The maximum number of classes was set to 40 and the maximum iteration set was 100.

The classification was done on time-series NDVI datasets for the years 2001, 2005, 2009, 2013, and 2017. The result for each are 40 ISODATA classes with no specific land cover class. Due to the coarse resolution of the data used, only Level 1 classification of the Anderson Classification Scheme was considered. These land cover classes are agricultural land, built-up, forest, rangeland, water, and wetland (Anderson, et al., 1976).

Assignment of ISODATA classes to each of the six land cover classes were done by NDVI profiling. Three sample pixels were randomly selected from each ISODATA class. Time-series NDVI values from Julian date-of-year 001 to 353 were determined for each sample pixel and the average value per date-of-year of the three samples was calculated and plotted. This process resulted in an NDVI profile for each ISODATA class. Ground truth data from Google Earth images were used to match an NDVI profile with a certain land cover.

2.5 Post-processing

A common issue with generating multitemporal land cover maps is the presence of illogical land cover transitions. An example of this is the change from agricultural land to built-up and then back to agricultural land. While this may not necessarily be impossible, this becomes less likely to happen for shorter time intervals. To address this, multivariate alteration detection (MAD) was implemented. This transformation algorithm is based on canonical correlation analysis (CCA) which investigates the relationship between two groups of variables such as spectral bands from two remote sensing datasets. This approach was applied in this study to detect significant changes between two NDVI datasets. Given two NDVI datasets expressed as vectors:

$$X = [X_1 \dots X_k]^T \text{ and } Y = [Y_1 \dots Y_k]^T \quad (2)$$

where k is the number of NDVI values in a dataset, the transformation change detection is given by:

$$X - Y = [X_1 - Y_1 \dots X_k - Y_k]^T \quad (3)$$

Optimizing the measure of change in a simple difference requires maximizing the variance which measures the quality of being divergent:

$$\text{Var}\{v_1(X_1 - Y_1) + \dots + v_k(X_k - Y_k)\} = \text{Var}\{v^T(X - Y)\} \quad (4)$$

High absolute values of $v^T(X - Y)$ indicate pixel locations with maximum change. Hotelling (1936) standardized the approach in determining the optimum value for v through CCA. In a standard CCA, linear transformation is performed on each set of variables resulting in linear combinations arranged by correlation or strength of similarity. The highest mutual correlation is called the first canonical variable, the next highest is the second canonical variable, and so on. Given the canonical variables U and V , and $\rho = \text{Corr}(U, V)$ is the correlation between them, the problem can be generalized by:

$$\sum XY \sum^{-1} Y Y \sum^T Y X a = \rho^2 \sum X X a \quad (5)$$

$$\sum^T Y X \sum^{-1} X X \sum X Y b = \rho^2 \sum X X b \quad (6)$$

where $\sum^{-1} X X$ and $\sum^{-1} Y Y$ are the covariance matrices of the two images, and $\sum X Y$ and $\sum^T Y X$ are the inter-image covariance matrices which are equal. The system of Equations (5) and (6) generates new images $U = [U_1 \dots U_N]^T$ and $V = [V_1 \dots V_N]^T$ which are called the canonical variates. The pair of elements (U_1, V_1) is the maximally correlated pair, and (U_2, V_2) is maximally correlated which is orthogonal to, or uncorrelated with, (U_1, V_1) . Pair differencing is performed after CCA generating a sequence of transformed difference images:

$$\begin{cases} U = \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \\ V = \begin{pmatrix} V_1 \\ \vdots \\ V_N \end{pmatrix} \end{cases} \rightarrow \text{MAD} = \begin{pmatrix} \text{MAD}_1 \\ \vdots \\ \text{MAD}_N \end{pmatrix} = \begin{pmatrix} U_1 - V_1 \\ \vdots \\ U_N - V_N \end{pmatrix} \quad (7)$$

Each element of the MAD image generated using Equation (7) is called a MAD variate, which can be used for change detection analysis. For the NDVI datasets used in this study, 23 MAD variates were generated ($N = 23$), each representing one date-of-year since the comparison was done between two successive years.

Statistical thresholding was done using the MAD variates of each pixel to identify whether it changed or not. MAD variates indicating significant change tend to deviate strongly from a multivariate normal distribution. Furthermore, assuming independence of orthogonal MAD variates, one may expect the sum of the square MAD variates after standardization to unit variance approximately follow the chi-squared (χ^2) distribution with N degrees of freedom. The MAD variate standardization function is given by:

$$z = \sum_{i=1}^N \left(\frac{\text{MAD}_i}{\sigma \text{MAD}_i} \right)^2 \quad (8)$$

Each pixel, having a corresponding value for z , can be given weights according to the:

$$P(\text{unchanged}) = 1 - P_{\chi^2, N}(z) \quad (9)$$

A small value of z typically implies greater probability that the pixel is unchanged. A statistical threshold was selected such that $P(\text{unchanged}) > t$, where t is the threshold value taken from the χ^2 distribution at $\alpha = 0.05$. The degree of freedom (df) was taken to be equal to N . Figure 2 shows an example of a χ^2 distribution plot. The broken line marks the threshold value determined using the χ^2 distribution table. A pixel whose z value is greater than, or lying to the right of, t is classified as unchanged. Otherwise, the pixel is marked as changed pixel.

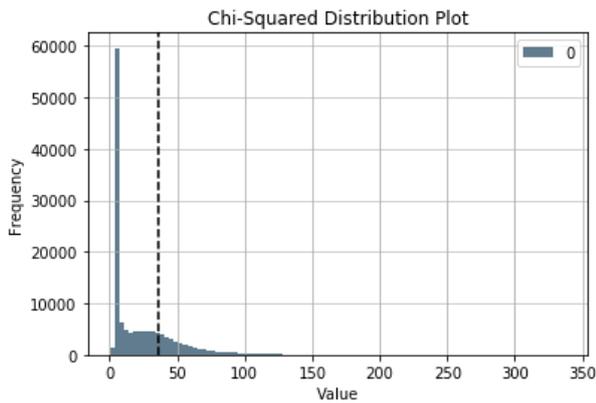


Figure 2. Sample χ^2 distribution plot

MAD was implemented to generate a mask for unchanged and a mask for changed land cover pixels based on NDVI values. Refinements in the initial land cover maps were done pairwise starting with a reference map (i.e., land cover at time t_0), which was the one with the highest overall accuracy. Land cover classification for unchanged pixels were retained. For changed pixels, the classification in the map being refined (i.e., land cover at time t_1) is adopted. The land cover at t_1 was then used as reference to refine the land cover map at time t_2 , and so on.

2.6 Accuracy Assessment

To generate a validation data, the watershed area was divided into 1 km by 1 km grids within each of which a validation point was chosen. The land cover of this point was noted based on ground truth data from Google Earth images. This process generated a total of 1,296 validation points used for checking the accuracy of each land cover map produced. A set of validation dataset was generated based on image data for each of the years 2005, 2010, and 2015 used for assessing the accuracy of 2001, 2005 and 2009, 2013, and 2017 land cover maps, respectively. The five-year interval follows the same period for updating official land cover and land use maps by the national mapping agency.

2.7 Vegetation Hot Spot Analysis

Vegetation hot spot mapping based on NDVI time-series data involved two stages: (i) creating space-time cube and (ii) emerging hot spot analysis. A space-time cube is three-dimensional representation of a time-series dataset whose base denote the spatial component or location (X, Y) and whose height indicate the temporal component or timestamp of the data (Kraak, 2003). Each element of a space-time cube is a bin, which represents a specific instance of the entire dataset at a given location and time. Bins that share the same location represents a bin time series, which is the dataset for that specific location through time. Bins with different locations but share the same temporal extent make up a time slice (Esri, n.d.). In the NDVI dataset, each bin time series corresponds to a unique pixel location while a time slice corresponds to a time-step interval of one year.

Emerging hot spot analysis uses the space-time cube as input data along with parameters such as neighborhood distance and neighborhood time step. The neighborhood distance indicates the spatial extent of the analysis to be considered for each location while the neighborhood time step gives the number of time-step interval to be included in the analysis. The analysis first

calculates the Getis-Ord G_i^* statistic for each bin to test whether there is significant clustering of high or low values at each location considering spatial neighbors. Each bin is given a hot spot classification based on the results. Mann-Kendall test is then performed to identify whether there is a significant trend in each bin time series to determine the final hot spot classification for a specific location (Esri, n.d.).

Annual average NDVI data from 2001 to 2017 were used as input to create the space-time cube. Inland water bodies such as lakes were masked out so that the analysis focused only on the watershed's land area. The NDVI raster data were converted to polygon grid features with the same dimensions as the raster pixels and with object ID, date, and NDVI as attributes. The object ID was set to be the same for a unique pixel location regardless of the year while the date was set to January 01 for all pixels belonging to the same year. NDVI values (range 0.1-0.8) were negated so that originally low values would be taken as greater in numerical value in the hot spot analysis. This was done because the statistical calculations tag locations where clustering of high values occur as hot spots. The negation ensured that locations of decreasing vegetation (low NDVI in the actual data) over time were correctly identified as hot spots.

Vegetation hot spot analysis was done using Esri's Space Time Mining Toolbox in ArcGIS Pro. Cold spots were not considered as these indicate locations of healthy and dense vegetation and are not a primary concern of this study. A spatial neighbor distance equal to one pixel and neighborhood time step of 2 were used as input parameters.

3. RESULTS

3.1 NDVI Profiles of Land Cover Classes

The NDVI profiles generated for each of the 40 ISODATA classes were grouped together based on similarity in trends and land cover class. Different trends were observed within the same land cover. This is due to the varying characteristics of the surface. Figure 3 shows the NDVI profiles of the ISODATA classes categorized under each land cover type.

Water 1, with NDVI value ranging from -0.1 to 0.0 , shows a stable trend throughout the year, representing water areas such as most portions of Laguna Lake. Water 2 has the same NDVI and trend as Water 1 during the first half of the year but with spikes towards the latter part. These spikes represent increases in NDVI due to relatively large areas covered by water hyacinths which accumulate for several days near aquaculture structures in the lake or close to the lake shore while moving very slowly around the lake. Water 3 represent areas very close to the shore where water hyacinths accumulate more frequently with longer retention period than they do in areas far from the shore, hence.

The profile for wetland areas show increasing NDVI during the dry season and decreasing as the wet season starts. These areas, which are located along the lake's shoreline, are lands that get exposed when the lake water level is lower than average during the dry season months. Patches of vegetation, particularly grasses, grow on the exposed land, which explains the NDVI values that are as high as 0.6 . As the water level increases again during the wet season, these lands get submerged again giving lower NDVI values similar to those of water.

All NDVI profiles for ISODATA classes identified as built-up areas show a similar trend that is relatively steady for the entire year, having only changes within ± 0.1 . These trends indicate

land cover that barely changes, which is true for built-up areas. The difference in NDVI values between different built-up ISODATA classes are influenced by the amount of sparse vegetation present. One class, for example, is characterized by areas dominated by built-up environment with more sparse vegetation than another class, which represents high density urban areas.

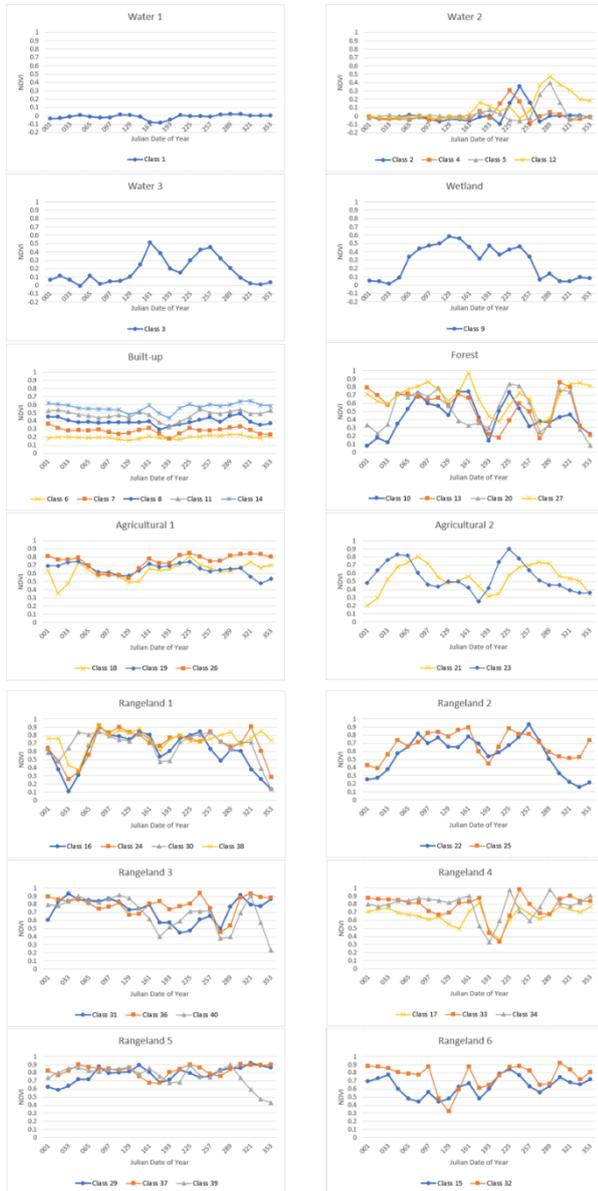


Figure 3. NDVI profiles of different land cover classes

Two types of agricultural land cover classes were identified based on their NDVI profiles. Agricultural 1 correspond to perennial cropland areas, which are for plants with longer agricultural cycles such as fruit trees. The profile shows a steadily high NDVI values which may drop after several months or even more than a year. Agricultural 2 shows the profile of croplands with agricultural cycles of five to six months as shown by two peaks within a year. One cycle shows clearly the phenology of the crops in which the increasing NDVI marks the growth period, the peak indicates the time when the crops are fully grown, and the decreasing NDVI happens during the harvest season. The minimum in the plot marks the post-harvest

period when the soil is almost bare until new crops are planted, and a new cycle starts. Agricultural 3 also represents croplands with two cycles. However, the period is likely to be more than six months and starts and ends in the middle of the year compared to Agricultural 2 whose first planting season starts in January.

The NDVI profiles of ISODATA classes identified as forest cover are more erratic. This behavior may be attributed to the fact that forested areas in Laguna Lake watershed are found in high altitude areas such as mountains. Because of this location and topography, a considerable portion of the forests are usually covered by thick clouds which are too difficult to be corrected even with the corrections applied on MODIS data and the SavGol filtering done on the time-series NDVI. However, image interpretation and validation showed that these classes belong to forestlands. In general, however, the four ISODATA classes show agreement in the trend during some portions of the year, such as high values during the dry season and drop in NDVI during some wet season months when persistent cloud cover is present.

Finally, six different profiles for rangeland areas were identified. The behavior of these different profiles may be attributed to the different phenological characteristics of grass and shrub species which cover the rangeland areas. For example, grassland areas are almost bare during the summer months when precipitation is low, while some types of bushes may shed leaves on certain seasons while other species may have abundant leaves throughout the year. Separation of specific vegetation types and plant species was not done in this study.

3.2 Multi-temporal Land Cover Mapping

Land cover maps were generated using the NDVI datasets for years 2001, 2005, 2009, 2013, and 2017. Table 1 shows that post-processed land cover maps showed improved overall accuracy and kappa coefficient compared to the initial maps. The map for 2005 was not post-processed because it was used as reference in the MAD implementation for change detection. Kappa coefficient values show moderate agreement with ground truth data (Congalton, 2001).

Year	Overall Accuracy		Kappa Coefficient	
	Initial	Post-processed	Initial	Post-processed
2001	0.75	0.79	0.57	0.64
2005	0.79	0.79	0.67	0.67
2009	0.69	0.72	0.52	0.55
2013	0.67	0.69	0.49	0.51
2017	0.69	0.71	0.51	0.54

Table 1. Accuracy assessment of MODIS NDVI-derived land cover maps

Table 2 shows the user's (UA) and producer's accuracies (PA) for each class in the post-processed land cover maps. Accuracy values range from 0.50 to 1.00, with the low values calculated for vegetation classes (agricultural land, forest, and rangeland) and the high values for the water and wetland classes. Confusion among vegetation classes are due to lack of more refined separation between the three vegetation land cover types as an effect of using a 250-m resolution data. Figure 4 shows the final land cover maps.

Class	2001		2005		2009		2013		2017	
	UA	PA								
Agricultural	0.63	0.63	0.59	0.66	0.53	0.67	0.50	0.54	0.51	0.62
Built-up	0.84	0.74	0.83	0.79	0.79	0.74	0.68	0.65	0.63	0.73
Forest	0.80	0.51	0.69	0.72	0.58	0.53	0.56	0.62	0.73	0.62
Rangeland	0.82	0.89	0.86	0.84	0.95	0.79	0.78	0.76	0.81	0.74
Water	0.71	0.88	1.00	0.88	0.95	0.88	1.00	0.89	0.95	0.77
Wetland	0.71	0.63	0.80	1.00	0.63	0.83	0.75	1.00	0.83	0.83

Table 2. Individual accuracies of different land cover classes

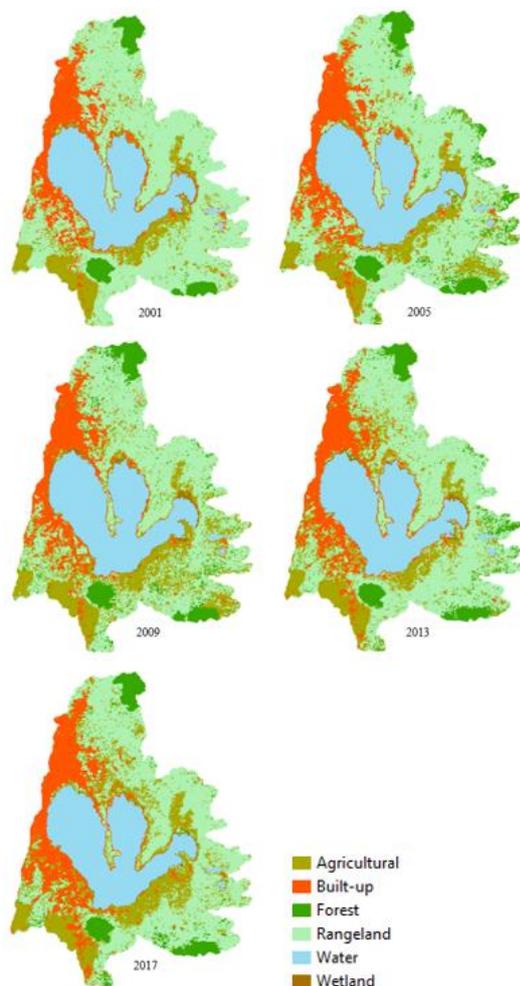


Figure 4. MODIS NDVI-derived land cover maps

Table 3 shows the percent area coverage of each land cover class per year. Rangeland, which includes grasslands, shrublands, and bushlands, covers most of the watershed at 39%-48%. This is followed by waterbodies, which includes Laguna Lake, at around 20%. Built-up environment makes up 12%-15% of land cover. Forested areas account for 4%-7% of the watershed, while wetlands cover 1% of the area. The total area of the watershed based on the land cover maps generated is 4,087.90 sq. km.

	2001	2005	2009	2013	2017
Agricultural	12.62	14.45	18.47	15.10	18.76
Built-up	12.43	13.29	12.22	13.15	15.43
Forest	4.22	7.35	6.67	6.72	5.64
Rangeland	48.65	43.58	40.02	43.09	39.25
Water	21.20	20.66	21.50	21.07	20.50
Wetland	0.88	0.68	1.12	0.86	0.42

Table 3. Percent area coverage per land cover class

3.3 Vegetation Hot Spot Analysis

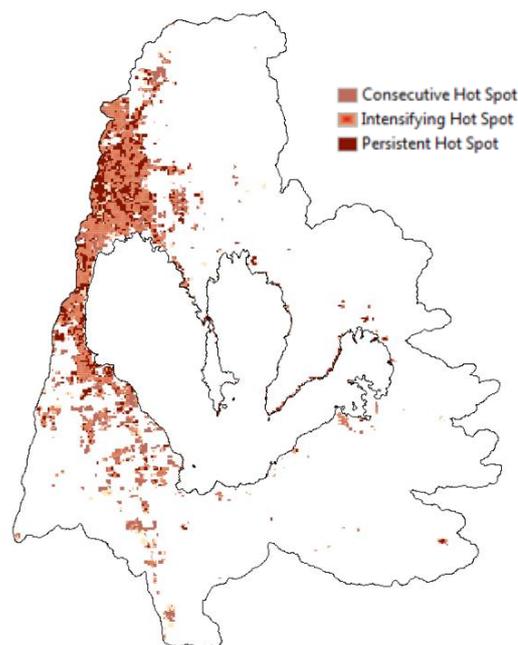


Figure 5. Vegetation hot spot map of Laguna Lake watershed

Figure 5 shows the distribution of vegetation hot spots based on the analysis of 2001-2017 NDVI data for Laguna Lake watershed. Hot spot areas cover 12.36% of the watershed. Of these hot spots, 39.55% are intensifying, 25.63% are persistent, and 20.70% are consecutive hot spots. Majority of the identified hot spots, particularly intensifying and persistent ones, are in the northwest side of the watershed, which is the high-density urban area of Metro Manila. Consecutive hot spots dominate the low-density urban towns of Laguna in the southwest. Hot spot areas can also be found along shoreline areas and rural town centers where clusters of built-up structures are found. These environments are characterized by sparse vegetation.

4. DISCUSSION

Analysis of the changes in built-up area coverage showed an overall increasing trend from 2001 to 2017, with as much 2.28% or 93 sq km increase in area happening between 2013 and 2017. Despite the very little fluctuations in area (less than 1% or 40 sq km), water and wetlands area coverage have been generally decreasing. This change may be attributed to land reclamation which has been one of the challenges faced by the lake's monitoring agency in managing the lake and watershed resources. Reclamation of land within the lake's shoreline area is allowed only after securing a clearance from the agency. However, some illegal reclamation activities have still occurred especially in the West Bay shoreline areas due to its proximity to highly urban Metro Manila. The reclaimed lands are commonly used for residential, commercial, and industrial purposes. Figure 6 shows the plot of built-up, water, and wetland area coverages from 2001 to 2017.

Figure 7 shows the percent area coverage of agricultural land, forest, and rangeland. agricultural land showed an overall decreasing trend while rangeland and forest cover were decreasing. One possible scenario is the conversion of rangelands to croplands to help in the demand for food supply in rural areas. In general, vegetation cover in Laguna Lake

watershed steadily declined from 65.49% to 64.91% (change of 0.11%-0.24% or 4-10 sq km) between 2001 and 2013. The decrease in area was highest, however, between 2013 and 2017, with as much as 1.27% or 51 sq km of vegetation loss. The increase in built-up area along with the decrease of vegetation cover can be explained by conversion of land to built-up environment. In the Philippines, a common example is the development of agricultural lands into mixed residential and commercial land use, which is the usual case for Cavite, Laguna, and Batangas. Industrial facilities are also common, such as in the case of western and southern Laguna where different industrial parks are located.

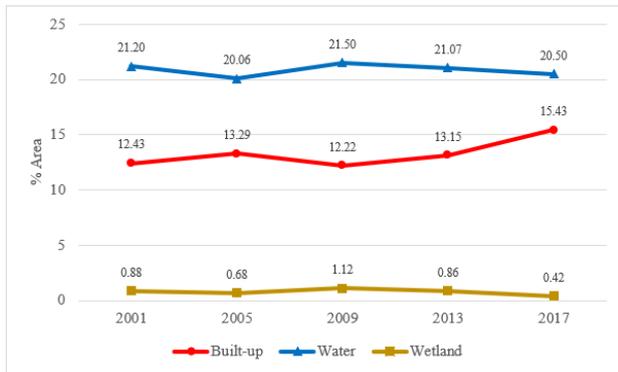


Figure 6. Percent area coverage of built-up, water, and wetland

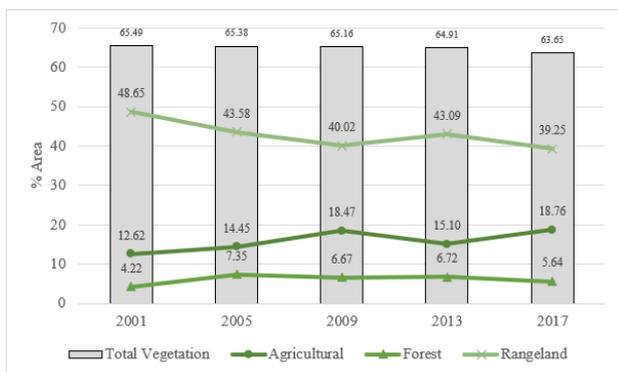


Figure 7. Percent area coverage of vegetation cover types

This finding is supported by the result of hot spot analysis. Intersecting the hot spot map with the vegetation cover per year revealed that 41.57% of the identified hot spots were vegetated areas in 2001. This decreased to 35.98% in 2005, 32.92% in 2009, 27.29% in 2013, and 19.27% in 2017. The decreasing trend confirmed that vegetation cover in hot spot areas were removed and the lands were converted to non-vegetated surface.

Persistent hot spots are characterized by long-term low to absent vegetation cover. These are found in high-density urban areas such as Metro Manila, in which built-up environment has been in place for at least 15 years. The presence of intensifying hot spots indicates already low but continuously declining vegetation cover, which happens in low-density urban areas experiencing continuous development, for the past 15 years including 2017. Consecutive hot spots appear to be areas of concern, as these are locations where low vegetation have been observed in the recent years. If declining vegetation persists, these hot spots may intensify.

5. CONCLUSIONS

The use of MODIS NDVI data to monitor land cover in a watershed such as the Laguna de Bay region provides an alternative to remote sensing data acquired by optical sensors with limited overpass period. Using time-series NDVI allows the identification of different land cover classes based only on statistical values and trends in the data and without the need for training dataset for classification. The result is a land cover map that is moderately acceptable with respect to ground truth data, which is enough for certain types and levels of applications.

Analysis of multi-temporal land cover maps of Laguna Lake watershed revealed changes in land cover types that are driven by anthropogenic actions as shown by the overall increase in built-up areas along with declining vegetation cover. This is also reflected in the observed decrease in vegetation cover in identified hot spot areas, indicating conversion to non-vegetated surface for the 17-year period studied.

The monitoring and assessment of land cover in Laguna Lake watershed is an important component in lake resource management. While findings shown in this study may be considered an initial assessment of the area, these can already provide insights on the current state of the watershed and on the steps that need to be taken not only by the lake's monitoring agency but by all stakeholders in order to mitigate the impacts of land cover in the lake water quality, which is the primary concern in Laguna Lake.

Although the use of moderate resolution data such as MODIS provided moderately acceptable maps, additional steps may be taken, however, to improve the results. This may include exploring the use of satellite data with higher spatial resolution but moderate temporal resolution and using ancillary data to improve the classification. Implementing other algorithms, such as Markov chain models, to address inconsistencies in multi-temporal land cover maps may also be considered. With these, more specific land cover classification may lead to better understanding of land cover transformation occurring in the area.

ACKNOWLEDGEMENTS

This research was done as part of the Multi-platform and Cross-sensor Water Quality Monitoring (MCWQM) Project. The Project was implemented by the University of the Philippines Training Center for Applied Geodesy and Photogrammetry (TCAGP), through the support of the Department of Science and Technology (DOST) of the Republic of the Philippines and the Philippine Council for Industry, Energy, and Emerging Technology Research and Development (PCIEERD). The Project was also done in collaboration with the Laguna Lake Development Authority (LLDA).

REFERENCES

- (n.d.). Retrieved October 25, 2019, from MODIS Land Team Validation: https://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD13&_ga=2.62252712.526663290.1572839604-536403554.1566955206
- (n.d.). Retrieved October 25, 2019, from MODIS Moderate Resolution Imaging Spectroradiometer: <https://modis.gsfc.nasa.gov/>

- Anderson, J. R., Hardy, E. T., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensing data*. United States Geological Survey. doi.org/10.3133/pp964
- Ball, G. H., & Hall, D. J. (1965). *ISODATA, A Novel Method of Data Analysis and Pattern Classification*. Menlo Park: Stanford Research Institute. Retrieved from <https://apps.dtic.mil/dtic/tr/fulltext/u2/699616.pdf>
- Carroll, M., Townsend, J., Hansen, M., DiMicelli, C., Sohlberg, R., & Wurster, K. (2010). MODIS Vegetative Cover Conversion and Vegetation Continuous Fields. In *Land Remote Sensing and Global Environmental Change* (p. 726). New York, New York: Springer. doi.org/10.1007/978-1-4419-6749-7_32
- Chen, J., Jonsson, P., Tamura, M., Gu, Z., Matsuhita, B., & Eklundh, L. (2004). A simple method for reconstructing high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, 322-344. doi.org/10.1016/j.rse.2004.03.014
- Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed spatial information. *International Journal of Wildland Fire*, 321-328. doi.org/10.1071/WF01031
- Dhodhi, M. K., Saghi, J. A., Ahmad, I., & Ul-Mustafa, R. (1999). D-ISODATA: A Distributed Algorithm for Unsupervised Classification of Remotely Sensed Data on Network of Workstations. *Journal of Parallel and Distributed Computing*, 280-301. doi.org/10.1006/jpdc.1999.1573
- Didan, K. (2015). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi.org/10.5067/MODIS/MOD13Q1.006
- Esri. (n.d.). *Create Space Time Cube From Defined Locations*. Retrieved from ArcGIS Pro: <https://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining>
- Esri. (n.d.). *How Emerging Hot Spot Analysis Works*. Retrieved from ArcGIS Pro: <https://pro.arcgis.com/en/pro-app/tool-reference/space-time-pattern-mining/learnmoreemerging.htm>
- Kraak, M. (2003). The Space-time Cube Revisited from a Geovisualization Perspective. *21st International Cartographic Conference*. 1988-1996.
- Laguna Lake Development Authority. (2017). *2017 Annual Report*.
- Nyamekye, C., Osei Jnr, E., Oseitutu, A. (2014). *Classification of time series NDVI for the assessment of land cover using NOAA/AVHRR data*. *Journal of Geomatics*, 34-39.
- Santos-Botja, A., Nepomuceno, D. S. (2003). Experience and Lessons Learned Brief for Laguna de Bay, Philippines. *Lake Basin Management Initiative ILEC/LakeNet Regional Workshop for Asia: Sharing Experience and Lessons Learned in Lake Basin Management*.
- Savitzky, A., & Golay, M. (1964, July). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 1627-1639. doi.org/10.1021/ac60214a047
- United States Geological Survey. (n.d.). doi.org/10.5067/MODIS/MOD13Q1.006