ASSESSMENT OF MODIS NDVI DATA AS ALTERNATIVE SOURCE OF INFORMATION FOR FOREST COVER MONITORING

Alita D. Sangalang, M. Sc.[1] and Ariel C. Blanco, Dr. Eng.[2]
[1]Resource Data Analysis Branch, NAMRIA, Fort Bonifacio, Makati City 1201
alitasangalang@yahoo.com
[2]Chairman, Department of Geodetic Engineering, UP, Diliman, Quezon City 1101
ayeh75@yahoo.com; ariel.blanco@up.edu.ph

ABSTRACT

Periodic observation of the country's forest cover is a must for proper planning and management of this important natural resource. Remote sensing technology has long been providing the world the means for mapping and monitoring natural resources for a variety of purposes. Notwithstanding the growing applications and benefits of this technology, the acquisition cost of remote sensing data is still the main concern that has to be addressed. In view of the characteristics of freely-available MODIS data, this study focused on the premise that this can be a potential source of information on forest condition. As vegetation reflects much radiation in the near-IR, Normalized Difference Vegetation Index (NDVI) can be used to identify vegetation and its broad classes. In this study, ISODATA technique was utilized in order to assess the usefulness of the Sav-Golay-filtered 250meter MODIS NDVI time series image covering Cagayan River Basin (CRB), Northern Luzon, Philippines. Through zonal analysis, mean NDVI Spectral Profile and amount of estimated per-cent cover from the Land Use/Land Cover Map (LULC) [13] were generated and the relationship was established to interpret the unknown classes. The final classified image showed significant improvement after classifying the misclassified pixels in the second iteration. The overall accuracy and kappa coefficient have increased by 9% and 0.1405 respectively. In addition, confidence interval for forest category obtained a very small amount of variance wherein at 90% confidence level, the confidence limits are 82.15% and 84.19%. Results showed that the method developed and described in this study can be used to extract forest land class from a 250 meter MODIS NDVI time series image and can therefore become valid as alternative source of information for forest cover monitoring.

Keywords: *Remote Sensing, MODIS Vegetation Indices, Zonal Analysis, ISODATA, Forest Land Class*

INTRODUCTION

Forest is considered one of the most valuable natural resource available to man. Various important functions of forest provide a great help in uplifting economic status of a developing country, among others, biodiversity conservation, ecological balance, environmental stability and food security. Other functions include regulating water, acting as carbon sink to lessen global warming and preventing soil erosion, landslide and flood.

In 1900, there were around 21 million hectares of Philippine forest out of the 30 million hectares of total land area [5]. In 2003 report, "The State of the World's Forests", the Philippines' rate of deforestation was at 1.4% annually from 1990 to 2000 [6]. Based on Philippine Forestry Statistics, our forests are indeed disappearing with about 7.2 million [7] hectares or only 24% of the country's land area as forested in 2003. The 2010 Land Cover Map which is the latest available data revealed 6.53 million hectares of forest [10]. This necessitates the use of certain measures to effectively monitor and manage the forest for sustainability and protection.

The advent of remote sensing technology provided the world the means for monitoring natural resource information on its condition for better management. Forest is a key application of remote sensing. For Philippine forest, aerial photographs were the basis for the first ever released Forest Resource Condition Map (FRCM) at a scale 1:50,000 [10]. The succeeding initiatives as regards to forest information utilized conglomeration of Landsat, SPOT, SPOT5 and AVNIR images [10]. Remote sensing images are mostly commercial data and would be very expensive to acquire since to monitor the entire country, say, annually, would require large data acquisition. It is also difficult to acquire images with minimal cloud cover.

Periodic observation of the country's forest cover is possible considering the technical and scientific aspect of MODIS, having near real-time, wider spatial resolution and high temporal resolution, this can really be a potential source of alternative information in monitoring changes in the country's forest cover. As vegetation reflects radiation in the near-IR, Normalized Difference Vegetation Index (NDVI) can be used to identify forest. The NDVI is related to the amount of greenness with values ranging from -1 to 1. One of the many data produced by MODIS is the MOD13Q1 250 m NDVI image derived from the first two bands of MODIS' 36 bands necessary to derive NDVI with Band 1 in visible red (0.62-0.67 micrometers) and band 2 in near infrared (0.841-0.876 micrometers). For this reason, there is a need to assess the usefulness of freely-available remote sensing datasets, with consideration to its spatial and temporal resolution at minimal cost by determining the feasibility of forest cover classification from a 250 meter resolution MODIS NDVI time series image to serve as alternative source of information for forest cover status.

MATERIALS AND METHODS

Study Area

Cagayan River Basin is located in the Northeastern part of the Island of Luzon. It is situated within 15° 45' and 18° 30' latitude and between 120° 45' and 122° 30' longitude. It is the country's longest and widest river basin. The land area is approximately 2,766,276 hectares [7]. The approximate perimeter of the basin is 882 kilometers [7] encompassing, fully or partly, the provinces of Cagayan, Isabela, Nueva Vizcaya, Quirino, Ifugao, Kalinga, Apayao, Mountain Province and Aurora.

Topography and Climate

The topography of the area is level to rolling (0-18% slope) at the central part of the basin including river banks that is around 38% of the area. Around 35% of the basin is mountainous owing to the mountain ranges that lie on both sides of the basin, i.e. Sierra Madre in the east and Cordillera Mountain Ranges in the west. The remaining portion of the basin is rolling to hilly. The elevation in the expanse of the study area is 0 to 2000m above sea level with some small portion at 2100-2800m which are peaks mostly at Cordillera Mountain Ranges.

Rainfall is deemed to be the most important element of climate in the country. Based on the Climate Map of the Philippines [20], most of the expanse of CRB falls on Type III climate where there is a short dry season, either during the months of March/April and November/December. The eastern part of the basin is on the east coast receiving heavy rainfall. Tropical storms hit the country between June and December and the northeastern part of the Island of Luzon is the most affected. These could be the reasons why cloudiness is experienced for longer period of time in a year. These have caused the images to be mostly cloud-covered.

Data Used

MOD13Q1: 16-day 250m NDVI is a higher level MODIS land product distributed as Hierarchical Data Format-Earth Observing System (HDF-EOS) projected to a tile-based sinusoidal grid. One tile unit is approximately 1200km by 1200 km in Julian date. It is one of the 6 MODIS Vegetation Index products generated from the daily MODIS Level-2G (L2G) surface reflectance, pointer file, geo-angle file and 1-km state file [15]. One hdf file of MOD13Q1 contains 4-band surface reflectance (blue, red, NIR & MIR), NDVI and EVI images.

2010 Land Cover Map of the Philippines [9] was generated by visual interpretation of 2010 AVNIR2 for the large expanse of CRB. The 2006-2007 SPOT5 and a single scene of 2010 Landsat image was used to fill some portions with no available 2010 AVNIR2 image. The visually-interpreted map was field validated to come up with the final Land Cover Map of the country for year 2010.

LULC Map [13] was generated from two sets of Landsat 7 TM and ETM+ dated 2009 to 2010 within the coverage of path 116, rows 47, 48 and 49. The 5 land cover classes contained in this map are forest, vegetation, water, bare soil and built-up areas that has obtained 91% overall accuracy and kappa coefficient of 0.88.

Cagayan River Basin Boundary is a polygon file in WGS 84 UTM Zone51 projection. This file was used to clip the various images and other related polygon files.

Methods

In producing forest cover map and in order to assess the usefulness of MODIS NDVI data, this study employed automated processing methodology through ENVI™, ArcMap™ and Quantum GIS™ (QGIS). The general procedure applied in the study is shown in Figure 1.

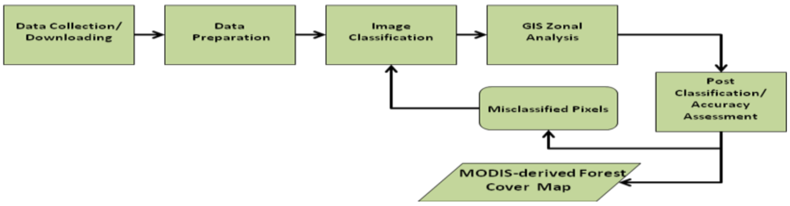


Figure 1. Generalized Flowchart of the Analytical Procedure

Data Collection/Downloading and Preparation

The MOD13Q1: 16-day 250-m VI products were searched and downloaded from the internet [17] via FTP protocol with inclusive dates of January 1, 2010 to December 19, 2010 for a total of 23 image files. Each of the multi-dated MOD13Q1 was reprojected to WGS84 UTM Zone51 and was subset to the boundary of CRB. A time series image was built by stacking the individual reprojected and subset MOD13Q1 images. The Sav-Golay Filter [4], a simple but robust method, was used to reduce/remove noise/contamination in the time series image using the optimal filtering window size (4, 6) [4] smoothening the data enough that the temporal detail of the time series image is still maintained. The 2010 Land Cover Map of the Philippines was collected from NAMRIA, a government agency that produced and published the said map. NAMRIA uses FAO classification system in land cover mapping. The map in vector file format was subset to the polygon file of the boundary of CRB and converted to raster file format. The land cover categories were generalized following the Anderson's scheme to conform to the MODIS NDVI time series-derived Forest Cover Map. Afterwards, the polygon file of CRB subset is converted to raster file in five land cover categories namely: (i) Agricultural land, (ii) Rangeland, (iii) Forest, (iv) Water, (v) Others. The LULC Map [13] was reclassified to produce a binary map for each land cover category. The reclassified images were used as input value raster in the GIS zonal analysis in order to calculate the percentage of each land cover category present in each of the unknown classes. The map was then downsampled to the MODIS NDVI pixel size and used as ground truth image.

Classification of MODIS NDVI Time Series Image using default parameters

The unsupervised ISODATA classification algorithm was applied to the time series image. This technique is able to distinguish unknown classes by clustering the pixels based on minimum spectral distance technique [2]. There was no prior knowledge on the existence or names of the classes [14], only a user-defined number of classes and inputs of some thresholds are needed and by some processes the classes should later be labeled. ISODATA was able to identify 20 spectral clusters as observed in the output single band raster layer.

GIS Zonal Analysis

Zonal analysis was done in order to give an indication as to the land cover classes in aid of labeling the 20 unknown classes of the MODIS NDVI classified image. Through zonal analysis, the relationship between (i) the derived mean NDVI spectral plot for each class and (ii) estimated percentage of area per land cover that is dominant in each class was analyzed. The NDVI spectral profile being one of the bases in identifying the unknown classes are derived from a multi-temporal image and was found to exhibit predictable variation [7].

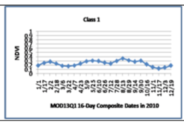
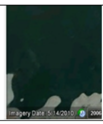
Computation of estimated area (%) of all the land cover types of the 20 classes

This was done to come up with the area in percent of all land cover types for each polygon and consequently for each class using the reclassified images as references. The land cover type with the highest computed area in percent was used as one indication to identify names of the 20 unknown classes, i.e., (i) Built-up, (ii) Forest, (iii) Soil, (iv) Vegetation or (v) Mask. In this procedure, reclassified values of LULC Map [13] were used as input value raster images while the polygon file of MODIS NDVI Classification image was the input vector zone data for all reclassified images.

Selection of Sampling Data from Google Earth™ Maps

Since 2005, Google Earth™ has provided maps by superimposing high resolution satellite images of the portions of earth surface available online with where one can view and make use of. The images are of varied resolution, dates and places of availability[18]. [2] and [11] have utilized images in Google Earth™ and have proven its worth as reference image and as substitute for in situ observation respectively. Depending on the places, the resolution can be as high as 0.1 m but generally, it is at 15 m resolution [18]. The available 2010 images were accessed through KML (Keyhole Mark-up Language) files of the individual polygon files of the twenty 20 unknown classes of MODIS NDVI classified image. The sample output of the three procedures to identify unknown classes was summarized in Table 1.

Table 1. Identification of unknown ISODATA class

MODIS NDVI Classified Image	Result of Zonal Analysis		Validation from Google Earth™		Final Land Cover Category and Remarks
	NDVI Spectral Profile	Estimated Cover (%)	Landcover	Coordinates	
Class1		95.92% Water		1860862.64 334087.18	Low NDVI values w/ high % water & as validated from Google Earth Class 1->**Water**

Post Classification/Accuracy Assessment

The 20 unknown classes were labeled and the classes that belong to same land cover category are combined to make it comparable with the categories present in the reference image. There were a total of 3 categories for the MODIS NDVI time series-derived Forest Cover Map I, i.e., forest, vegetation and water and was compared with LULC Map [13].

Classification of MODIS NDVI image using input parameters

MODIS NDVI time series image was classified again following the previous procedure but using the computed class standard deviation and distance as input parameters such that ISODATA can better identify unknown classes. The trends in the NDVI spectral profile that was identified previously will serve as reference profile. This was done to depict land cover characteristics better and determine if the unknown classes generated by ISODATA can be identified by interpretation of the variation in the NDVI spectral profile even in the absence of high resolution satellite images.

Through zonal analysis, NDVI spectral profile was generated, examined and characterized of its trends that were attributed to land cover. The unknown classes were labeled according to the interpreted mean NDVI Spectral Profile of each of the classes based on the Anderson's scheme [1]. The output is MODIS NDVI time series-derived Forest Cover Map II and was compared to the 2010 Land Cover Map. Based on the Producer's and User's Error, there are misclassified pixels that merit further consideration. Although, efficient techniques are applied in the process, the classified images are often not without errors [9]. Despite the fact that the classification of MODIS NDVI time series obtained considerable accuracy in reference to its spatial resolution and that misclassification errors are inherent once the image is divided into different land cover classes, finding and classifying again the misclassified pixels may prove that these pixels actually belong to certain land cover class. Improving the classification accuracy by dealing with misclassified pixels are subject of studies by [16] and [12].

The ISODATA technique was performed with the misclassified pixels extracted from the MODIS NDVI time series-derived Forest Cover Map II. Class standard deviation and distance were also computed and used as input parameters. The unknown classes were labeled and same classes were combined accordingly. The process was done twice. All the correctly classified pixels and the remaining misclassified pixels were merged that resulted to final MODIS NDVI time series-derived Forest Cover Map II and this image was compared with the 2010 Land Cover Map of the Philippines.

Confidence Interval

In this study, a confidence interval was computed only for the Forest land category. Ten thousand random samples of 250m pixels were arbitrarily chosen from the 2010 Land Cover Map where the final MODIS NDVI time series-derived Forest Cover Map II was compared. Figures in the confusion matrices of final MODIS NDVI time series-derived Forest Cover Map II compared to the both of 2010 Land Cover Map and random samples were used in the construction of confidence interval. An estimate of an unbiased variance [19] was done following Equation (1) while the confidence interval [19] at 90% confidence level was computed following Equation (2).

$$Var(\hat{p}) = \frac{N-n}{N}\frac{\hat{p}(1-\hat{p})}{n-1} \quad (1) \qquad \hat{p} \pm t_{1-0.9/2,n-1}\sqrt{\mathrm{var}(\hat{p})} \quad (2)$$

RESULTS AND DISCUSSION

Table 2. Confusion Matrix for ISODATA Classification using default parameters

Classified Image	Reference Data	Producers Accuracy for FOREST (%)	User's Accuracy for FOREST (%)	User's Accuracy for 3 classes (%)	Overall Accuracy (%)	Kappa Coefficient
MODIS NDVI time series-derived Forest Cover Map I	Land Use/Land Cover Map (Principe, 2012)	86.44	84.31	78.36-84.31	81.88	0.6349

Forest Cover Map I obtained 86.44% (Table 2) accuracy for forest, a feature of primary interest in the classification. The range of user's accuracy (78.36%-84.31%) means that the map with 3 classes is reliable. It is 82% accurate when compared to LULC Map [13], and is reasonably good and acceptable considering the spatial resolution of 250m data and the classes that were discriminated. The Kappa Coefficient of 0.6349 indicated that the information in the classification image is in moderate agreement with the ground truth image. The image used in the classification is a one-year data in 250 meter resolution while the reference image was derived from a conglomeration of single-dated 30 meter resolution Landsat [13]. The LULC Map [13] that obtained a high accuracy of 91% is still not in perfect agreement. Relatively, the accuracy of the classified image being compared cannot be expected to exceed the accuracy of the reference image and this means that the accuracy of 82% is acceptable.

The mean NDVI spectral plots were utilized here by observing its patterns where the distinction of classes have been associated with land cover classes in the classified MODIS NDVI time series image. Similar pattern in the NDVI spectral profile can be attributed to pixels with mixture of both classes or with their growth pattern due to changing seasons. The existence of mixed pixel have been noticed just like in any other images and it is certain, considering the spatial scale of the NDVI image especially in the transition of the land cover classes.

Table 3. Classification Accuracy for ISODATA technique using input parameters

Classified Image	Reference Data	Producers Accuracy for FOREST (%)	User's Accuracy for FOREST (%)	User's Accuracy for 4 classes (%)	Overall Accuracy (%)	Kappa Coefficient
MODIS NDVI time series-derived Forest Cover Map II	2010 Land Cover Map of the Philippines	85.88	77.48	53.63-77.48	70.06	0.5526
MODIS NDVI time series-derived Forest Cover Map II (final)		87.70	83.46	71.12-83.46	79.44	0.6931

For Forest Cover Map II, forest land category has the highest percentage detection wherein around 77% (Table 3) of all forest pixels are actually forest on the ground, considering that there are 4 land cover classes that were identified as compared to the 3 classes of the first output. Because the more classes that are classified, the more confusing it is for the classifier. But the trends distinct to forest land from other profile have caused its highest accuracy among the 4 land cover classes.

The final Forest Cover Map II revealed a significant improvement. The class accuracy of forest land has improved by around 6% and remains as having the highest detection of the 4 land cover classes. The overall accuracy and Kappa coefficient has increased from 70% to 79.44% and 0.5526 to 0.6931 respectively, and that is something which is related to the study done by [16] which found out that misclassified pixels are often one of the main influences on final classification accuracy.

The computation of confidence interval obtained a very small amount of variance and showed that the accuracy of the forest land category, when the map was compared with the CRB subset of

2010 Land Cover Map [10] and with the random samples were not very different. Results showed that at 90% confidence level the true class accuracy for forest land lies between 82.15% and 84.19%.

CONCLUSION

This study revealed that the use of freely available MODIS Vegetation Indices data proved to be of advantage due to high temporal and wide spatial resolution, low cost of processing and storage requirement due to low volume of data and demonstrated the potential of time series derived therefrom for forest cover classification. The method developed in this study showed that forest cover information can be extracted from a 250 m MODIS NDVI data time series image through ISODATA technique. Results have indicated that acceptable levels of forest classification accuracy were achieved and it can be concluded that MODIS NDVI is a good alternative source of information for forest cover monitoring. The classification approach employed in this study is very promising and can be replicated to cover the whole country considering the availability of data at minimal cost. Results of this study can be integrated with other data as indicative map and used as input to various projects of concerned government entities that are needed by policy-makers and decision-makers in planning and management of natural resources.

ACKNOWLEDGMENT

The authors would like to thank NAMRIA, thru the Administrator, Dr. Peter Nilo Tiangco for equipments and data provided during the course of the study and DENR for the financial support.

REFERENCES

[1] Anderson, J.R., et al., (1976). A land use and land cover classification for use with remote sensing data, U.S.G.S. Professional Paper 964, Washington, U.S. Government Printing Office.
[2] Campbell, J. B. (2006). Introduction to remote sensing (4th ed.). The Guilford Press; 4th edition.
[3] Cha, S. & Park, C., (2007). The utilization of Google Earth images as reference data for the multi-temporal land cover classification with MODIS data of North Korea. Korean Journal of Remote Sensing, Vol.23, No.5.
[4] Chen, J., Jonsson, P., Tamura, M., Gu, M., Matsushita, B. & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time series set based on the Savitzky-Golay filter.
[5] ESSC. 1999a. The decline of the Philippine forest. ESSC, Ateneo de Manila University, Q. C.
[6] FAO, 2003.State of the World's Forests, Rome.
[7] FMB/DENR, 2011. Philippine forestry statistics. DENR, Visayas Avenue, Quezon City.
[8] Knight, J., Lunetta, R., Ediriwickrema, J. and Khorram, S. 2006. Regional Scale Land-Cover Characterization using MODIS-NDVI 250 m Multi-Temporal Imagery: A Phenology-based Approach**.** *GIScience and Remote Sensing*.
[9] Landgrebe, D. (1999) "Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data," Information Processing for Remote Sensing, chapter one, Edited by C. H. Ghen, World Scientific Publishing.
[10] National Mapping and Resource Information Authority (NAMRIA), Makati City.
[11] Niemisto, A, Lukin, V., Shmuleich, O., Yli-Harja, O. and Dolia, A., (2001). A Training-based Optimization Framework for Misclassification Correction.
[12] Perera, K., Herath, S., Apan, A. and Samarakoon, L. (2010) Application of MODIS 250 m images without in situ observations for mapping Mekong River Basin land cover. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XXXVIII, Part 8, Kyoto Japan 2010430.
[13] Principe, J. (2012). Exploring climate change on watershed sediments yield and land cover-based mitigation measures using SWAT Model, RS and GIS: case of Cagayan River Basin, Philippines.
[14] Richards, J. & Jia, X. (1999). R. S. digital image analysis, 3rd ed. Berlin: Springer-Verlag.
[15] Solano, R. Kamel Didan, Jacobson, A. and Huete, A., 2010. MODIS Vegetation Indices (MOD13) C5 User's Guide, Verlag.
[16] Wang, Y. & Civco, D. Post-Classification of Misclassified Pixels by Evidential reasoning: A GIS Approach for Improving Classification Accuracy of Remote Sensing Data.

[17] http://edcsns17.cr.usgs/newearthexplorer (accessed on August 2011)
[18] http://en.wikipedia.org (accessed on September 2012)
[19] http://www.yale.edu/ceo/OEFS/Accuracy_Assessment.pdf (accessed on 14th March 2013)
[20] PAG-ASA. http://kidlat.pagasa.dost.gov.ph/ (accessed July 2010)