Asociación Española de Teledetección

Revista de Teledetección 35, 94-108
ISSN: 1988-8740

Caso práctico

A MODIS generated land cover mapping of Honduras: a base-line layout to create a national monitoring center

S. Rivera[1]*, J. L. Lowry[1], A. J. Hernández[1], R. D. Ramsey[1], R. Lezama[2] and M. Velázquez[2]

[1] *Remote Sensing and GIS Laboratories. Department of Wildland Resources. College of Natural Resources. Utah State University. Logan. Utah 84322-5230. USA*

[2] *Geographic Information Center. Department of Forestry Applied Research. National School of Forest Sciences-ESNACIFOR Siguatepeque. Honduras*

Abstract

The Remote Sensing and GIS Laboratory at Utah State University (USU) began the process of establishing a remote sensing national land cover monitoring center in order to monitor land cover and use in Honduras. A national land cover map derived from MODIS (Moderate Resolution Imaging Spectrometer) imagery products was developed. We designed a protocol for interpretation and analysis of MODIS data products that included a Google-Earth on-screen sampling scheme, a field data collection of training samples and a classification tree algorithm. The first land cover map prototypes and algorithms were developed using a time series of MODIS 2007 and 2008 imagery, elevation data (STRM) and a time series of MODIS's Enhanced Vegetation Index (EVI). In the model validation, the Kappa coefficient was K = 65.1% and the overall model accuracy was 70%. This map will serve as a base line to monitor future land cover changes in Honduras.

Key words: mapping, MODIS, tropics, Google, EVI.

Resumen

Generación de un mapa de cobertura del suelo en Honduras a partir de datos MODIS: una base para el diseño de un centro nacional de seguimiento

El Laboratorio de Teledetección y SIG en la Universidad Estatal de Utah (USU) inició el proceso de establecer un centro de monitoreo de detección de la cobertura y uso de la tierra a fin de monitorear la cubertura en Honduras. Un mapa de la cubertura derivado de MODIS (Espectrómetro de Imágenes de Resolución Moderada) fue desarrollado. Se diseñó un protocolo para la interpretación y análisis de productos de datos MODIS que incluyó un protocolo de Google-Earth para toma de muestras en pantalla, una colección de campos de datos de muestras de entrenamiento y una clasificación basada en un algoritmo de árboles de decisión. El primer prototipo de mapas y algoritmos fueron desarrollados usando una serie temporal de imágenes MODIS del 2007 y 2008, datos de elevación (STRM) y una serie de tiempo del Índice Mejorado de Vegetación de MODIS (EVI). En la validación del modelo, el coeficiente Kappa fue de K = 65,1% y la exactitud global del modelo fue del 70%. Este mapa servirá como línea base para monitorear cambios en la cobertura futura de la tierra en Honduras.

Palabras clave: Mapeo, MODIS, Trópicos, Google, EVI.

* Corresponding author: samuel.rivera@aggiemail.usu.edu

Received: 11-04-110; Accepted: 10-05-11.

Introduction

Land resources throughout the world are rapidly being depleted, and developing countries in the tropics are experiencing some of the most detrimental effects from rapid land use conversion. Some 10.4 millions of tropical forests were lost in the last five years (FAO, 2005). Central America and Mexico have the second highest global deforestation rate (Eggen-McIntosh, 1994). All Central American countries are classified with a *high threat* to loose their forest resources (Global Forest Watch, 2010). The global Forest Resources Assessment indicates that Honduras lost some 186,000 hectares of forest between 2000-2005, the highest annual deforestation rate in Central America: 2.8% (FAO, 2005). This causes an enormous environmental disruption and jeopardizes the weak economy and the well-being of the country.

Appropriate monitoring is an important step to determine how much forest resources should be harvested versus what should be conserved. Remote sensing approaches for assessing and monitoring forest resources provide a cost-effective means by which forest inventories and land use monitoring can be achieved. Various methods may be used to map vegetation patterns on the landscape, the appropriate method depending on the scale and scope of the project (DeFries *et al.,* 2004). Projects focusing on smaller regions, such as national parks, may rely on aerial photo interpretation. Mapping vegetation over larger regions has commonly been done using digital imagery obtained from satellites, and may be referred to as *land cover mapping*.

In 1995 the country of Honduras produced a land cover map with the help of the German government (COHDEFOR, 1996). This forest resources map contained 8 land cover classes (dense conifer forest, sparse conifer forest, mangrove forest, broadleaf forest, mixed forest, water bodies, neighboring country, non-forestlands) and was produced using a visual interpretation method of Landsat imagery. The methodology to produce the map was time-consuming and not easily replicable, thus making it a useful product at a single moment in time, but not useful for monitoring purposes. Currently it is outdated.

Some other attempts have been done in the last decade. Two land cover maps were developed. In 2001, a land cover map was developed by a World Bank sponsored initiative (PMDN) and another one by the Tropical Agricultural Centre for Higher Education and Research, CATIE (Ordonez and House, 2002). None of them had a transparent and replicable methodology and the government was still uncertain about how much area is covered by forest or other lands. In 2005, a forest inventory was conducted throughout the country. Around 300 sampling plots were established systematically all over the country. Due to the lack of funding the inventory was not complemented with the remote sensing assessment and rapidly became obsolete (AFE-COHDEFOR, 2006).

Our proposal is to produce a similar land cover map with the same 8 land cover classes (and possibly more) using contemporary MODIS satellite imagery in a fashion that can be done relatively rapidly and over time-specifically for long term monitoring purposes. This was not possible in 1995, and is indeed feasible today because of advancements in image processing technology since 1995 and the availability of MODIS imagery since 2000.

Utah State University, located in Logan, Utah had the responsibility of coordinating the development of field data collection protocols, mapping methodologies, and coordinating land cover mapping for the region. ESNACIFOR was responsible for field data collection, image and ancillary data preparation, and land cover modeling. The development and refinement of the legend was coordinated by ESNACIFOR with input from managers, government officials and investigators in Honduras. This paper presents an overview of the methodology used to create the national land cover dataset, and highlights several of the issues associated with achieving this product through a coordinated process between the two locations-countries.

The goal of this project was to develop and transfer a methodology (protocol) using MODIS (Moderate Resolution Imaging Spectrometer) imagery to create a national-level land cover map. Other products, such as a national-level ecosystems map —not described on this paper—, were derived from the land cover map. It should be emphasized that developing the map was not the sole goal of the project. In addition, the other goal of the project was to train Honduran scientists in the development of this map, so that they can continue monitoring national resources in the future.

Objectives

Specific objectives for this project were:

1) Develop an image classification methodology using MODIS imagery to create a national-level land cover map for the country of Honduras.

2) Train Honduran scientists in fundamental knowledge of remote sensing technologies, and remote sensing-based mapping.

3) Provide training to the scientists on the GIS methodologies (protocol) developed by USU for creating the national land cover map.

4) Demonstrate how the protocol can be used for future land resource monitoring as well as assessment of land cover change between 2009 and the future.

Project organization

Project study area

Honduras is located in the center of the Central American isthmus, between 13° and 16° latitude North and 83° and 89.5° longitude West (Fig. 1). It has an area of 112,088 square kilometers. It is a rich country in terms of natural resources and has the highest percentage of forest lands among the other Central American nations (AFE-COHDEFOR, 2006). Approximately 50% of the country is still covered by undisturbed forests (Richards, 1996) which include humid tropical forests, arid or deciduous tropical forests, cloud forests, mangrove wetlands, and pine forests. Forests in Honduras are being depleted at an accelerated rate. The deforestation rate is currently 80,000 hectares per year, which is one of the highest deforestation rates in the hemisphere (Stonich, 1993).

Tropical forests are typically associated with coastal mountains receiving high amounts of precipitation while pine forests are located in the headwaters of rivers in the mountains of central Honduras. In Honduras, two major river systems drain the central highlands to both the Caribbean Sea (eight river basins) and the Pacific Ocean (two river basins) (Gutiérrez, 1992; Laboranti, 1982). The average precipitation rate is 2,000 mm per year and this rainfall produces significant runoff from watersheds (Hargreaves, 1992).

The country is composed of extremely fragile ecosystems. Since it is a narrow strip of land, rivers run from the continental divide (2,000 meters above sea level) to the lowlands

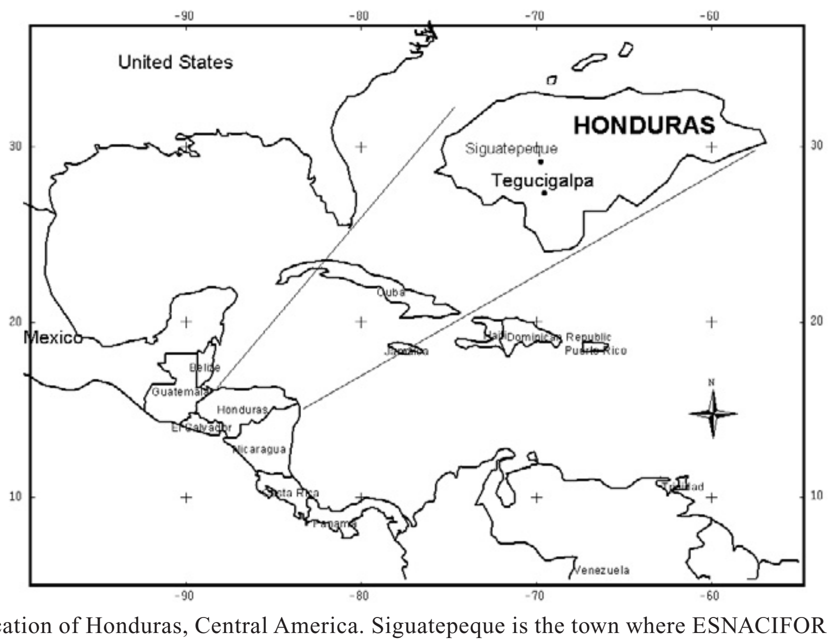

Figure 1. Location of Honduras, Central America. Siguatepeque is the town where ESNACIFOR is located.

in the Pacific and Caribbean coasts in very short distances. As a result, rivers are typically steep, enclosed in v-shaped valleys, and exhibit dendritic drainage patterns. Greater than half of Honduras might experience 300 mm of rain in 24 hours (Hargreaves, 1992). Soils are formed from metamorphic, volcanic, and sedimentary parent materials. As a result, eight of the ten world soil orders exist in Honduras. These soils are classified with a high to very high erosion risk. In addition, the terrain in Honduras is characteristically steep with 75% of the territory having slopes greater than 30% (Hargreaves, 1992). Forest cover removal produces an enormous environmental disruption.

Division of responsibilities

Overall project tracking and management was conducted by the RS/GIS Lab at Utah State University. The lab was responsible for developing the methodology and providing training to ESNACIFOR personnel in the use of the aforementioned methodology. ESNACIFOR provided the appropriate funding for the project, including salaries, fees, and travel expenses for all project participants of both parties. ESNACIFOR was also be responsible for providing access to possible field data from forest management plans from ICF: the Honduran Forest Service.

Project coordination and timeframe

Training workshop and the assessment of the ESNACIFOR's GIS lab capability were conducted in spring of 2008. Initial field data collection protocols were established by the lab at Utah State University. Google sampling data collection primarily occurred at the end of 2008 in Honduras. MODIS image processing and classification workshop dedicated to ensure consistent mapping methods was conducted during the winter of 2009 at Utah State University lab. Periodic meetings, teleconferences and field visits (in Honduras) were conducted in the summer of 2009 to ensure the collaborative mapping process. Additional field data to improve the model and for validation was collected from May through September of 2009. Mapping efforts were completed by the end of 2009. Last training session and completion of final products was held at Utah State University lab in April of 2010. The land cover map was completed and delivered to the public in May 2010.

Methods

Image selection and preparation

An important part of this project involved determining the best image classification approach for the monitoring protocol. The MODIS sensor provides an ideal remotely sensed platform for developing a national-level land cover/resource monitoring program (Muchoney *et al*., 2000; Loveland *et al*., 2000; Brown *et al.,* 2007). With 500 meter pixel resolution it is considered a moderate resolution imaging sensor and cannot be used for highly detailed land cover mapping (by comparison the Landsat sensor offers 30 meter pixel resolution). However, MODIS had several key advantages that made it highly suitable for this project. With 7 spectral bands, MODIS provides adequate spectral resolution to map vegetation. These 7 bands approximate the 6 spectral bands offered by the Landsat sensor.

1) The MODIS sensor orbits the earth each day (Vermote, 2008). The daily capture of MODIS imagery means MODIS can provide daily images of the Earth's surface, making it ideal for monitoring purposes.

2) In addition to the daily MODIS surface reflectance product, MODIS provides a suite of derived products. One of these products is the *Daily Surface Reflectance Quality Product* which is an 8-day composite of images intended to provide the best image pixels for an 8-day period (Vermote, 2008). This is helpful in Honduras where cloud cover presents a significant challenge for remote sensing-based mapping, as the best cloud free pixels are used in the composite image (Vermote, 2008).

3) The MODIS image scene encompasses the entire country of Honduras with a single swath. This is important because other sensors, such as Landsat, would require mosaicking

multiple image scenes (approximately 12 Landsat scenes are required to cover the country of Honduras) from several swaths, imposing considerable technical challenges.

4) Imagery and products derived from MODIS are offered without cost from the United States Geological Survey (USGS) and be downloaded from the World Wide Web (www.glovis.gov).

Generally, land cover mapping is accomplished by segmenting the landscape into areas of relative homogeneity that correspond to land cover classes from an adopted or developed land cover legend. Technical methods to partition the landscape using digital imagery-based methods vary. Unsupervised approaches involve computer-assisted delineation of homogeneity in the imagery and ancillary data, followed by the analyst assigning land cover labels to the homogenous clusters of pixels. Supervised approaches utilize representative samples of each land cover class to partition the imagery and ancillary data into clusters of pixels representing each land cover class. An important part of mapping involved the use of training data for image classification, whether a supervised or unsupervised approached was used. Sampling can be a time consuming and costly part of any remote sensing-based land cover mapping project. We determined the best approach to collecting sampling data through the course of the project. Possible options included using existing forest plot sampling sites (AFE-CODHDEFOR, 2006), collecting additional sampling plots, and most likely, a combination of both.

Land use classification system

Utah State University conducted some preliminary tests using MODIS imagery to create a national-level map of Honduras. These preliminary results were presented to the government authorities in spring of 2008. The preliminary results were developed through a lattice of points that were used to train the classification tree algorithm. These training points were labeled by drilling them through the 1995 forest resources map. The training samples were then imported to the classification tree algorithm and rules were generated to derive the resulting land cover map. The classes that were best mapped were: broadleaf forest, mangrove forest, and pine forest.

We decided to use a modified land use classification system used by the International Geosphere-Biosphere Programme (IGBP) (Jensen 2005). This classification system was selected because it has been used by the Honduran government authorities. This classification system has also been largely used by other projects either globally (Friedl *et al*., 2002; Loveland e*t al.,* 2000) or in the Central and South American Regions (Muchoney *et al*., 2000; Latifovica *et al.,* 2000; Brown *et al*., 2007), due to its capacity to fit the MODIS resolution.

In the final land use and cover map, 13 classes were delineated (Table 1). Initially eight classes were assumed to be mapped (Dense and sparse pine forest, broadleaf forest, mixed forest, urban, mangrove forest, water bodies and agriculture farms/pasture), however, through the whole process of Google-Earth sampling plus the field visits, five more classes were added to the classification (Commercial agriculture, shrublands, dry forest, savannas, and shrimp farms).

High resolution imagery sampling

Image interpretation was conducted using Landsat ETM with ArcGIS© and high resolution orthophotography using Google Earth™. Table 1 identifies which classes (10 classes initially) that we anticipated that were going to be identifiable with each method (UNESCO, 1983). Training data were obtained from three sources: Existing data, Image Interpretation of Landsat ETM and Orthophotography, and field samples. Existing data were available form two sources: Forest management plans, and point data collected by the FAO (AFE-COHDEFOR, 2006). These data were carefully checked to make sure that they were reliable, and then they were formatted in preparation for digital image classification. Formatting was done within a spread-sheet (e.g. Excel).

The sampling approach can be described as a systematic-selective hybrid sampling design. The image interpretation of Landsat ETM in-

Tabla 1. Legend used in the land cover mapping of Honduras

Label	Symbol	Classification System (adapted from IGBP Jensen, 2005
Dense Evergreen Needleleaf forest	BCD	Lands dominated by trees with a percent canopy cover >60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
Sparse Evergreen Needleleaf forest	BCR	Lands dominated by trees with a percent canopy cover between 30-60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
Broadleaf Forest	BLF	Lands dominated by trees with a percent canopy cover > 60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
Mixed Forest	BMX	Lands dominated by trees with a percent canopy cover > 60% and height exceeding 2 meters. Consists of tree communities with interspersed mixtures or mosaics of the other four forest cover types. None of the forest types exceeds 60% of landscape.
Shrublands	MAT	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover is > 60%. The shrub foliage can be either evergreen or deciduous.
Mangrove	BMG	Lands with a permanent mixture of water and herbaceous or woody vegetation that cover extensive areas. The vegetation can be present in either salt, brackish, or fresh water.
Water Bodies	LYL	Oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt water bodies
Agriculture and Pasture	AGP	Lands covered with temporary crops followed by harvest and a bare soil period (e.g., single and multiple cropping systems. It also includes natural or planted pasture for livestock.
Commercial Agriculture	AGC	Land covered by perennial crops such as bananas, pineapple, and oil palm.
Urban	URB	Land covered by buildings and other man-made structures. Note that this class will not be mapped from the AVHRR imagery but will be developed from the populated places layer that is part of the Digital Chart of the World.
Woody Savannas	SAB	S Lands with herbaceous and other understory systems, and with forest canopy cover between 30-60%.The forest cover height exceeds 2 meters.
Dry Forest	BSE	Lands dominated by trees with a percent canopy cover > 60% and height exceeding 2 meters. Consists of seasonal broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods. Dominant species are: *Simarouba glauca, Switenia humilis, Casia grandis, Mimosa sp., Albizia guachepele, Sterculia apétala, Enterolobium cyclocarpum, Karwinskia calderonii, Crecentia alata, Bursera simaoruba, Leucaena salvadorensis, Tabebuia rosea, Gliricidia sepium, Lysiloma sp., Bombacopsis quinata* and *Cedrela odorata.*
Shrimp Farms	CAM	Formerly covered by Mangrove forest and now they are ponds-water bodies dedicated to shrimp production. All sites are located in the Pacific coast at an elevation below 200 meters above sea level.

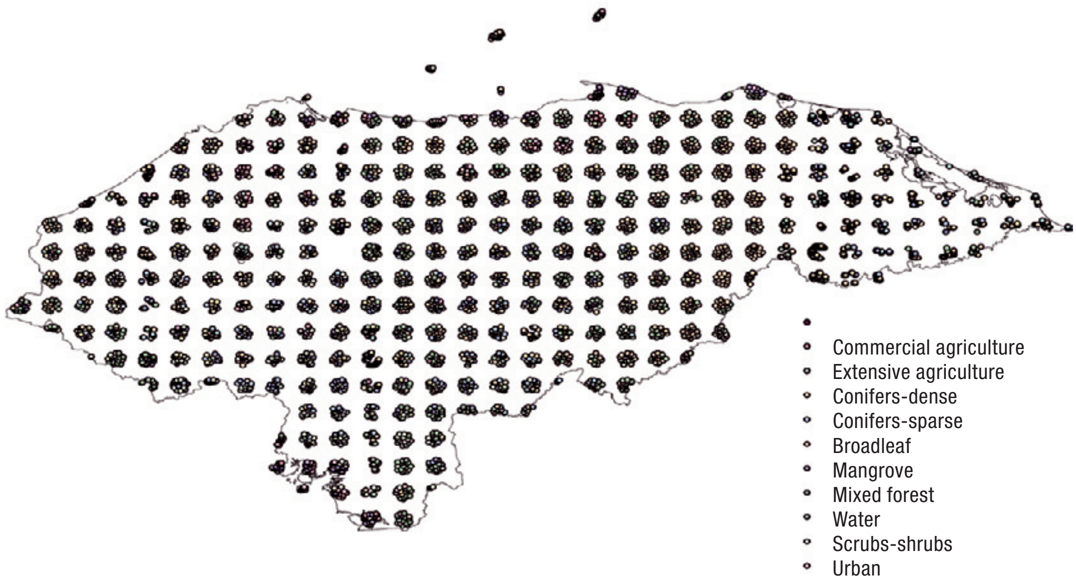

Figure 2. Site areas within which point locations were opportunistically chosen to be used as training data. These are 311 locations containing more than 5,000 points that were photo interpreted in Google Earth images.

volved the assigning label attribute to each point location on a systematic grid of site areas. There were 311 site areas on a grid covering the entire country (Fig. 2). Each site area represents an area of about 80 km^2 (10 km diameter). Within each site area the image analyst chose the best point location representative of the land cover classes in the site area. Between 10 and 15 point location samples were chosen for each site area, and samples were at least 1 sample point apart. The analyst tried to get samples for as many different land cover classes seen in the site area polygon.

Google-Earth™ data collection protocol

Google Earth as a scientific and environmental visualization tool has been increasingly used since its launch in 2005 (Sheppard and Cizek, 2009). We basically used the same procedure that was used with the interpretation of Orthophotography in Google Earth™. The site area and point location feature classes had to be converted to .kml format and were used with Google Earth™ to interpret higher resolution imagery (though perhaps not as up-to-date as the ETM imagery). The best way to use Google Earth™ was to add the site area and point location .kml files to Google Earth™, and use them as a point of reference. The analyst had the option to place the road network or other layers in Google Earth© for reference. With ArcMap© open on one screen and Google Earth™ open on another screen, the analyst used Google Earth™ to interpret the imagery, but as added the label in ArcMap©. Some 5,616 samples were collected to build the training data set.

Ground truth data collection

Training and map validation data were collected through ground-based field work to supplement the existing field and Google Earth™ collected data (5,616 data points). The ground-based field samples were 240 samples, collected by traversing navigable roads in three different a mapping zones (North coast, South and Western part of the country) and opportunistically selecting plots that met criteria of appropriate size (500 by 500 mts or 25-hectare minimum) and composition (stand homogeneity). Field data were collected using ocular estimates of biotic and abiotic land cover components, including percent cover of dominant classes. Laptop computers, Landsat imagery, digital orthophotoquads, and GPS devices were used for navigation and plot identification whenever possible. Each plot was identified with a UTM coordinate using a GPS. Field data were recorded onto paper field forms and subsequently entered into a database.

Predictor layers

Utah State University has considerable experience using a wide variety of image classification approaches. Most recently we completed a five year project mapping land cover over a large portion of the southwestern United States (Lowry *et al.,* 2007) using a classification tree algorithm. Several predicting layers were used in the primary classification. Spatial data layers used to map land cover included image-derived and ancillary datasets. Core image-derived datasets included individual seven bands of MODIS images from 2007 and 2008, the Enhanced Vegetation Index (EVI). Ancillary datasets were derived from 30 meter digital elevation models (DEM) obtained from the NASA's STRM (Shuttle Radar Topography Mission) (http://www2.jpl.nasa.gov/srtm/).

Modeling approaches and procedure

The USU-ESNACIFOR team investigated several avenues for image classification. In particular we experimented with methods similar to those used in previous large landscape mapping efforts such as the 1995 Utah GAP land cover project (Homer *et al*., 1997). The cluster-busting method was the first modeling approach that we used. In this method we visually determined an appropriate number of clusters. Then we ran an initial map with isodata, to create a cluster map with maximum likelihood classifier using the signature file from isodata. Then we used this output to subsequent steps.

The next method was the use of decision tree classifiers that are well suited for land cover mapping. Classification and regression trees (CART) were developed by Breiman *et al.* (1984) and were quickly recognized as a valuable tool for discriminating complex relationships among environmental variables (Friedl *et al.,* 2002). Decision trees readily accept a variety of measurement scales in addition to categorical variables, and have demonstrated improved accuracies over the use of traditional classifiers (Hansen *et al*., 1996; Pal and Mather, 2003). Finally, decision tree software is readily available, computationally efficient, and by using a hierarchical approach to define decision rules, is intuitive to a variety of users.

For our project we incorporated the decision tree software See5 (RuleQuest Research, 2004) with ERDAS Imagine®. The tool, developed for the National Land-Cover Dataset 2001 (Homer *et al.,* 2004) project (hereafter «NLCD mapping tool») provided the ideal solution to our need for an efficient integration of the decision tree software within a spatially explicit modeling environment.

Using the NLCD mapping tool, decision tree models were generated in See5 (RuleQuest Research, 2004) with the boosting option, and then spatially applied in ERDAS Imagine®. Modeling was iterative and subsequent iterations tested using different combinations of predictor datasets, or additional samples in an attempt to improve the model. An iterative process of adding/subtracting predictive layers from the model produced, finally, a more refined map.

Map validation using Google Earth™

Sample polygons were generated from the final land cover map. First, the raster version of the map was converted to a vector polygon dataset. The size of land cover polygons, for all classes, in the final land cover map was heavily skewed toward many small polygons and a few large polygons (i.e. a Poisson distribution). Randomly selecting a set of polygons from such a distribution results in many small polygons and very few large polygons (by simple probability given such a distribution). For our selection process we wanted to randomly select polygons that were large enough that an assessment of their accuracy could tell us something about the map at the stand level. We also wanted to select polygons that were within the high resolution portion of Google Earth™ Images. In other words, we wanted to identify polygons where land cover could be interpreted from higher resolution imagery, and which were not too small.

To begin, we recognized that all the land cover classes do not have the same spatial pattern; which is often a result of their unique ecological characteristics and abundance on the landscape. Some classes may be considered «matrix» communities and are spatially represented by large contiguous areas. Other classes are best described as «patch» communities that are imbedded within the broader matrix. Based on the notion of matrix and patch communities (Poiani *et al.,* 2000) we divided the mosaic of land cover classes into two groups based on their proportional abundance on the landscape. Figure 3 shows the division between matrix and patch classes. Examples of large matrix classes include *Broadleaf Forest*, *Dense Pine Forest*, and *Sparse Pine Forest* while examples of patch classes are *Mixed Broadleaf-Pine Forest*, *Dry Forest*, and *Urban Areas*.

Candidate polygons for accuracy assessment were first selected based on whether they were within the high resolution portion of Google Earth™ Imagery (Fig. 4). Then they were selected based on size. Matrix classes were selected as potential sample polygons if they were between 1,000 hectares and 2,000,000 hectares in size. Patch classes were selected if they were larger than 100 hectares, the minimum mapping unit for the map. Following these two selecting rules up to 60 polygons were randomly selected for each land cover class. Table 1 identifies the number of randomly selected polygons for each class, the total area sampled (sum of sample polygons) and the percent of the total map the sampled portion represents. A total of 685 polygons were chosen as sample polygons to be used for accuracy assessment. Figure 4 shows the selected polygons (Table 2). The process basically consisted on the analyst using numeric codes for label assignment. The Google™ file (.kml) was later converted to ArcGIS© file (shapefile).

Results

Land cover map

The most important and identifiable mapped classes were Broadleaf forest and pine forest (sparse and dense) that cover more than half of the country (58.1%) (Table 3, Fig. 5). These forest types are located following the mountain chains that go from west to east and north to south in the central part of the country. They are distributed along this rough topography in which 70% of the territory is located over 30% slopes. The broadleaf forest is located along the north coast and in a corridor that goes from the middle to the north-western part of the country, following areas of higher precipitation. Pine forests are located in drier areas in smaller patches scattered all over the central and western

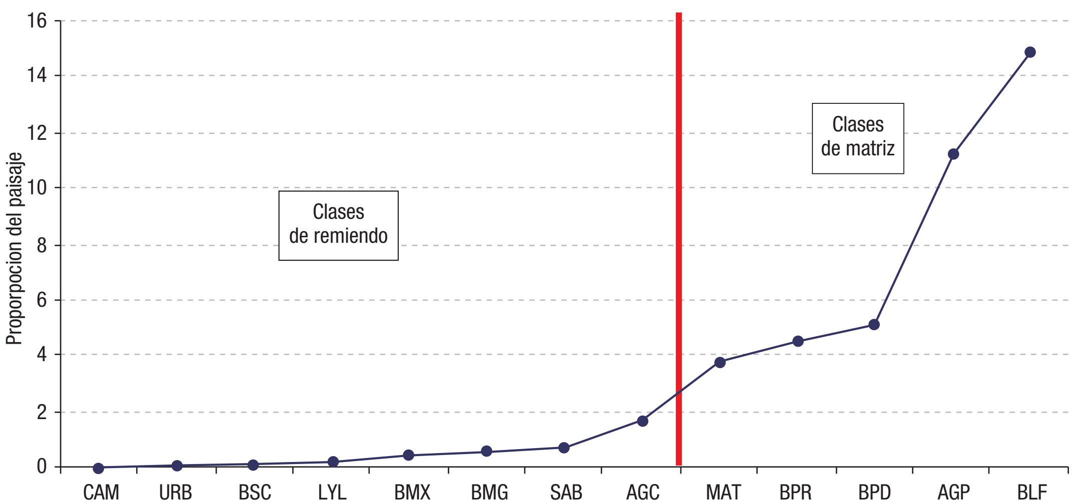

Figure 3. Matrix and patch land cover classes based on proportional abundance: basis for the map validation protocol using the Google-Earth sampling scheme.

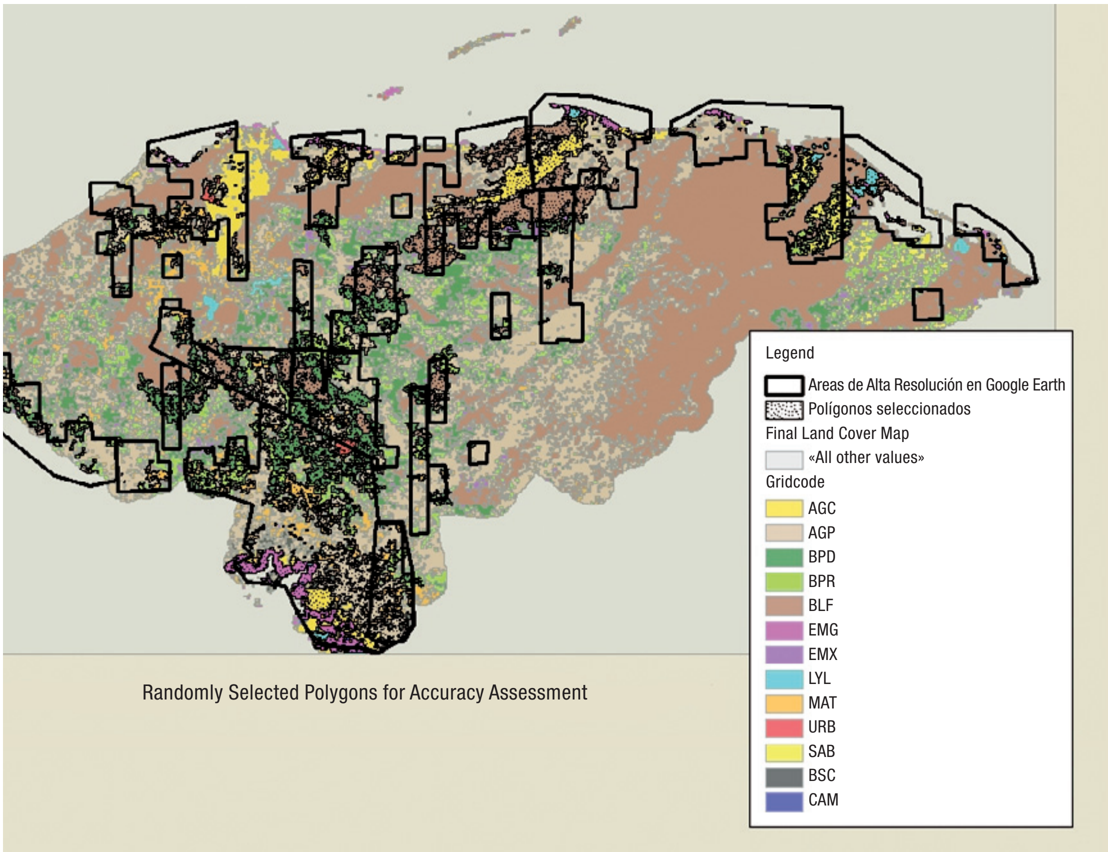

Figure 4. Randomly selected polygons for accuracy assessment and map validation.

part of the country. In between these patches, subsistence agriculture and pasture lands areas are found, forming a mosaic of agricultural and forest patches. In bigger patches, agro-commercial activities are located in the intermountain larger valleys. These areas are dedicated to high yield crops such as: bananas, cantaloupes, oil palms, citrics, sugar cane, and pineapples. Mangrove forest was found only in the pacific coast where most of this coastal ecosystem is located. Significant areas of mangrove forest have been cut and inundated to establish shrimp

Tabla 2. Randomly selected polygons to be used as reference data for the map validation procedure

Class name	Code	Numeric code	Type	Count	Sampled area (ha)	Total area (ha)	Percent area sampled
Ag. Farm/Pasture	AGP	2	Matrix	60	416,250	3,192,275	13.04
Dense Pine Forest	BPD	3	Matrix	60	437,375	1,448,275	30.20
Sparse Pine Forest	BPR	4	Matrix	60	227,600	1,273,625	17.87
Broadleaf Forest	BLF	5	Matrix	48	475,625	4,226,675	11.25
Shrubland	MAT	9	Matrix	60	202,375	1,088,825	18.59
Ag. Commercial	AGC	1	Patch	60	210,275	478,050	43.99
Mangrove Forest	BMG	6	Patch	60	147,100	192,875	76.27
Mixed Conf.-Broadleaf	BMX	7	Patch	60	16,475	136,775	12.05
Lakes & Lagoons	LYL	8	Patch	60	39,875	77,425	51.50
Urban Areas	URB	10	Patch	23	23,425	25,925	90.36
Savannah	SAB	11	Patch	41	95,550	218,500	43.73
Dry Forest	BSC	12	Patch	60	24,575	56,425	43.55
Shrimp Farms	CAM	13	Patch	10	2,325	3,600	64.58
				662	2,318,825	12,419,250	18.67

Tabla 3. Areas of land use and cover as a result of the land cover mapping of Honduras

Legend (name-code)	Area in 2009	
	km²	%
Commercial agriculture-AGC	4,042	3.63
Agriculture pasture-AGP	28,961	25.98
Dense pine forest-BPD	13,859	12.43
Sparse pine forest-BPR	11,919	10.69
Broadleaf forest-BLF	39,037	35.02
Mangrove forest-BMG	1,067	0.96
Mixed forest-BMX	673	0.60
Waters bodies-LYL	689	0.62
Shrublands-MAT	8,592	7.71
Urban-URB	204	0.18
Woody savanna-SAB	2,121	1.90
Dry forest-BSC	212	0.19
Shrimp farms-CAM	93	0.08
Total	111,468	100.00

farm ponds. We were able to identity these large areas. Dry forest was located in the southern portion of the country and in small areas in the central and northern regions where it is known that precipitation is limited in amount and period of occurrence, meaning less than 1,000 mm/year distributed in 3-4 months. Shrublands were also located mostly in the western and central regions and are often seen as areas of transitions where the forest has been cleared and/or burned and most of the cleared areas have remained untouched to produce a secondary forest. This is true for most of the cleared and/or burned areas for either pine or broadleaf forest. The urban class identified only when urban settlements were larger than 200,000 people. Smaller urban centers were hardly captured given the spatial resolution and spectral of the MODIS signal. The mixed forest was detected in the transitional zones between the pine forest and broadleaf forest. Due to its foliar composition, its detection was not very accurate. An explanation to this is discussed later in this paper. Honduras has only 2 important lakes: one is natural and the other is a man-made water body. They both are located in the central part of the country and were accurately mapped. The last class detected by the MODIS sensor was

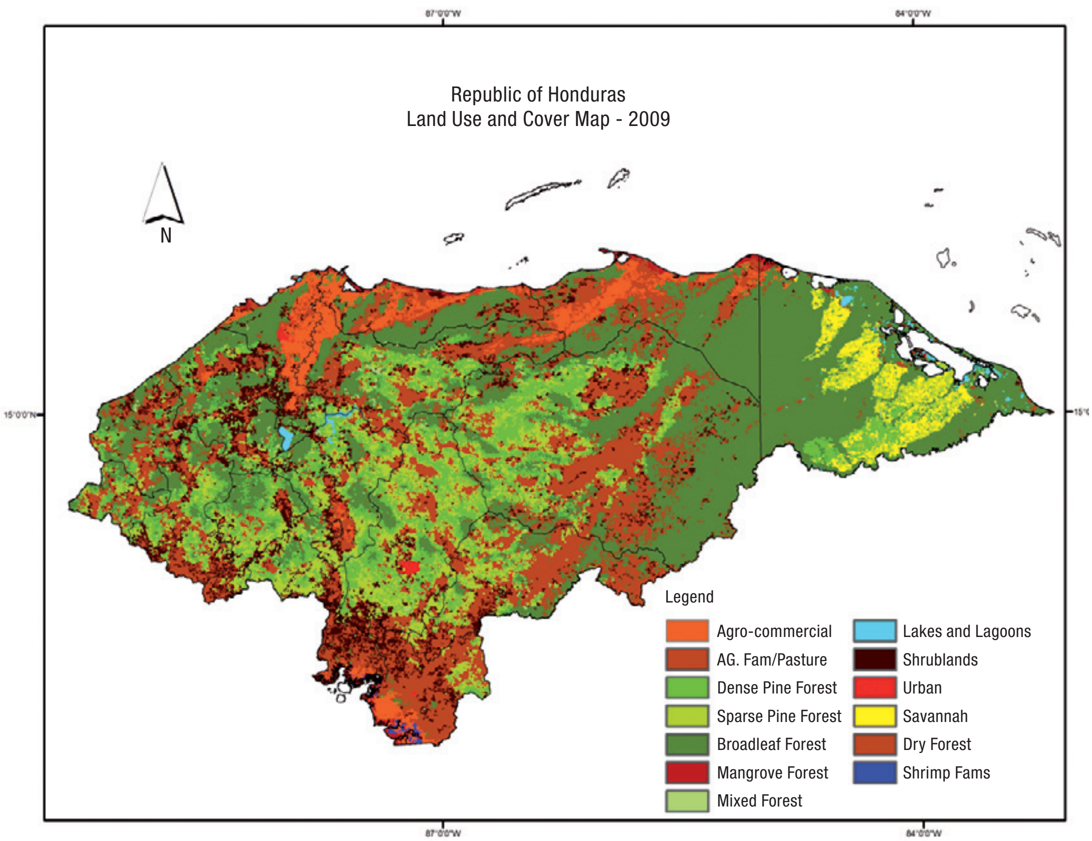

Figure 5. The MODIS imagery-derived map of land use and cover of Honduras, Central America.

the woody savanna. A peculiar ecosystem located in the northern and eastern portion of the country, close to the Caribbean coast. This is typically composed by grass and shrub vegetation accompanied by sparse pine trees. These areas are for high precipitation (between 3-4 meters per year), giving as a result highly leached soils, and sparse taller vegetation.

Model validation

The model validation has the purpose of measuring, as objectively as possible, the accuracy of the map. Usually this is given by providing a quantitative measure of map accuracy (Congalton and Green, 1999). The accuracy assessment was run using the Kappa Tool. The results showed that the Kappa Index was: 0.651258 (Standard error of kappa: 0.0055672, Z-Score for kappa: 116.981) see Table 4. The Kappa statistic can be interpreted as follows: Values below 0.40 would suggest the agreement between reference data and the mapped data is poor and could occur by chance. Values between 0.40 and 0.80 represents moderate agreement and values over 0.80 represents strong agreement (Congalton and Green, 1999).

The Kappa statistic for the land cover map was in the «moderate agreement» ranging indicating the results of the error matrix good. The Z-score for Kappa tells us whether the agreement between the mapped data and the reference data could occur by chance. A Z-score higher than 1.96 suggests that, at a 95% confidence level, the results of the error matrix do not represent chance agreement. The Z-scores for both maps are well above 1.96 indicating that we can have a high level of confidence in the error matrix and Kappa statistic.

The error matrix is also very useful to visually and quantitatively see which classes are highly confused with one another. This is done by examining the numbers in the off-diagonal cells. The error matrix for the land cover map also tells us that the overall accuracy of the map was 70%. However, the error matrix is most useful for telling us something about the accuracies of the individual land cover classes. The error matrix provides information on two types of error: 1) errors of commission, and 2) errors of omission (Jensen 2005). Errors of commission represent reference locations that were incorrectly mapped as other mapped classes, and are presented as the percentages at the right of the row totals (Table 4). For example, there were 415 reference samples that «landed on» AGC, but only 352 of those were AGC. The 90 samples that were not AGC, but landed on AGC are considered errors of commission. Errors of omission represent locations on the map that were not mapped correctly, and are presented as the percentages at the bottom of the column totals. For example, there were a

Tabla 4. Accuracy assessment results of the land cover mapping of Honduras

Class	AGC	AGP	BPD	BPR	BLF	BMG	BMX	LYL	MAT	URB	SAB	BSC	CAM	Row total	Row %
AGC	**352**	21	0	0	6	1	0	5	21	2	0	0	7	415	85%
AGP	0	**662**	0	7	7	0	0	0	99	10	0	0	0	785	84%
BPD	0	15	**583**	193	7	0	6	0	16	0	0	0	0	820	71%
BPR	0	103	31	**217**	0	0	10	0	24	0	57	0	0	442	49%
BLF	0	66	71	8	**627**	3	0	0	80	0	31	0	0	886	71%
BMG	1	23	0	3	0	**240**	0	10	9	2	10	1	0	299	80%
BMX	1	15	13	5	6	2	**4**	0	13	0	6	0	0	65	6%
LYL	4	19	0	0	2	5	0	**67**	4	0	3	1	1	106	63%
MAT	3	128	0	22	0	0	56	4	**160**	0	0	3	0	376	43%
URB	3	0	0	0	0	0	0	3	1	**47**	0	0	0	54	87%
SAB	0	0	0	5	0	1	1	0	0	0	**186**	0	0	193	96%
BSC	3	29	0	0	1	0	0	0	35	1	4	**6**	0	79	8%
CAM	0	0	0	0	0	0	0	5	0	0	0	0	**7**	12	58%
Column total	367	1,081	698	460	656	252	77	94	462	62	297	11	15	**4,532**	
Accuracy	96%	61%	84%	47%	96%	95%	5%	71%	35%	76%	63%	55%	47%		**70%**

Kappa: 0.651258. Standard error of kappa: 0.0055672. Z-Score for kappa: 116.9.

total of 367 AGC reference locations, but only 352 (96%) were in locations mapped as AGC. The remaining 4% were mapped as something else and are considered errors of omission.

Examining the error matrices for the map we can see that some land cover classes were mapped very well, while others were mapped quite poorly. If we consider any class mapped with accuracy greater than 80% mapped well, we note that AGC, BPD, BLF, and BMG are all mapped very well. Classes mapped moderately well (50%-80%) are AGP, LYL, URB, and SAB. Classes mapped poorly (<50%) are BPR, BMX, MAT, and CAM. For example in the map, we noted that 56 BMX reference locations were erroneously mapped as MAT. We can say that BMX was «confused» with MAT.

Discussion

About the modeling approach

A primary objective of this land cover mapping project was to develop a transparent methodology that was repeatable and could be consistently and independently applied by the Honduran scientists. The decision tree classifier combined with the Google™-sampling protocol met this objective well. We found the decision tree classifier considerably more time-efficient. The decision tree classifiers are a more powerful tool for discriminating land cover classes. Our results were very consisted with other investigations (Lativofic *et al.*, 2000; Friedl *et al.,* 2002). They are also a very interpretable method and explicit method, because their hierarchical decision rules and splits can be revealed and explained.

About Google Earth™ sampling and validation

Throughout the course of the project we recognized the importance of providing a measure of map quality to users of the land cover map. While limitations of time, money and logistics prohibited a formal accuracy assessment (i.e. external validation with probability-based sample design), we believe the methods we employed provide useful information to map users. Google Earth™ is a readily available tool that can be accessible to almost anyone in the world (Sheppard and Cizek ,2009). The novelty of this approach allows enormous saving in time and resources and training. We found very little references that confirm that this tool has been used to collect training samples in a land use classification map. We hope it can be used in the future.

Examining the error matrices for the map we can see that some land cover classes were mapped very well, while others were mapped quite poorly. If we consider any class mapped with accuracy greater than 80% mapped well, we note that AGC, BPD, BLF, and BMG are all mapped very well. Classes mapped moderately well (50%-80%) are AGP, LYL, URB, and SAB. Classes mapped poorly (<50%) are BPR, BMX, MAT, and CAM. The Google Earth™ validation sampling scheme proved to be a very cost-effective procedure. No references were found on the use of this approach for validation a land map use.

About MODIS Selection

MODIS capability and the classification procedure exceeded the project expected goals. Initially, 8 classes were set up in such way that can potentially be discriminated and mapped. As the project progressed, five (5) more classes were added to the classification algorithm. MODIS imagery proved to be very affordable and successful by identifying and discriminating land use classes at the country level. The accuracy assessment of the map was very high particularly for the some classes such as broadleaf forest, mangrove forest and commercial agriculture.

We also recognize that MODIS data had some limitations, especially at discrimination some classes such as: Sparse Pine Forest (BPR), Mixed Forest (BMX), and Shrublands (MAT). This is basically a limitation of the sensor to identify clearly these classes, which are transitional classes, at a resolution of 500 meters. However, these classes are considered a transition of secondary forest which is constantly growing to become primary forest or fully grown forest. In the case of Shrublands, they will grow until they become adult trees. Similarly, in the mixed forest, trees will mature until either conifer or broadleaf species become dominant. Therefore, we believe that the MODIS signal does not work very well at determing different vegetation stages at this resolution. Although it does very well differentiating fully grown or developed vegetation classes and types. Other researcher found these MODIS limitation in the past (Muchoney *et al.*, 2000; Latifovic *et al.*, 2000). Similarly Eggen-McIntosh *et al.* (1994) found the same obstacles using AVHRR imagery.

Summary

MODIS capability and the classification procedure exceeded the project expected goals. Initially, 8 classes were set up in such way that can potentially be discriminated and mapped. As the project progressed, five (5) more classes were added to the classification algorithm. MODIS imagery proved to be very affordable and successful by identifying and discriminating land use classes at the country level. The accuracy assessment of the map was very high particularly for the some classes such as broadleaf forest, mangrove forest and commercial agriculture.

The Google™ sampling protocol showed a high confidence as was used to collect the sampling training data. With minor training sessions, analysts were able to operate and collect samples. Despite the tedious work involved in the sampling collection, very few mistakes were made when interpreting the images. This protocol seems very promising as a cost-effective method to collect training samples without needing costly field visits.

The objective of the accuracy assessment was to quantitatively measure the accuracy of a MODIS derived land cover map product. The accuracy assessment described in this document outlines a methodology using freely available imagery through Google Earth™ to interpret reference polygons that were subsequently used to generate sample locations used in the accuracy assessment. Thirteen land cover classes were mapped and their agreement assessed with an error matrix and Kappa statistics. Overall accuracy was 70% for the generalized map product. This is comparable to other large landscape mapping efforts (Laba *et al.*, 2002; Lotsch *et al.*, 2003). Looking into individual classes through the error matrix reveals that some classes were mapped better than others. Fully grown vegetation classes such as AGC, BPD, BLF, and BMG are all mapped very well. Transitional vegetation classes, such as BMX were very poorly mapped and is highly confused with MAT, BPD, and BPR.

Acknowledgements

We gratefully acknowledge the financial support of the World Bank and the Government of Honduras, without which completion of this project could not have been possible. We also would like to thanks ESNACIFOR officials: Asdrubal Calderon, Gabriel Barahona and Daniel Galeano. And finally, we would like to recognize the vision to develop the land cover map of Honduras: Mr. Atilio Ortiz who passed away during the course of this project. His leadership and initiative made this work possible.

References

AFE-COHDEFOR. 2006. *Evaluación Nacional Forestal de Honduras 2005-2006*. Proyecto Apoyo al Inventario y Evaluación de Bosques y Arboles TC-HON-3001 (A). Resultados del Inventario de Bosques y Arboles 2005-2006. Honduras. Octubre 2006. 127 pp.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. & STONE, C. J.1984. *Classification and Regression Trees*. Belmont, CA: Wadsworth, Inc.

BROWN, J. C., JKASTENS, J., WARDLOW, B., JEPSON, W., CAMARGO COUTINHO, A., VENTURIERI, A., LOMAS, J. & K. PRICE, K. 2007. Using MODIS to detect cropping frequency variation in mechanized agriculture in Amazonia. *Anais XIII Simpósio Brasileiro de Sensoriamento Remoto*, Florianópolis, Brasil, 21-26 abril 2007, INPE, pp. 99-101.

COHDEFOR. 1996. *Plan de acción forestal de largo plazo 1996-2015*. PLANFOR. Borrador. Unidad de Planificacion Institucional. AFE-COHDEFOR. Tegucigalpa MDC, Honduras. 127 pp.

CONGALTON, R. G. & GREEN, K. 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton: Lewis Publishers.

DEFRIES, R. S., FOLEY, J. A. & ASNER, G. P. 2004. Land-use choices: balancing human needs and ecosystem function. *Front. Ecol. Environment,* 2(5): 249-257.

EGGEN-MCINTOSH, LANNON, S. K. B. & JACOBS, D. M. 1994. Mapping distributions of central America and mexico. USDA-USFS. Southern Experiment Station. New Orleans LA. USA. Proceeding reprint.

FALZARANO, S., THOMAS, K. & LOWRY, J. 2005. Using decision tree modeling in GAP analysis land cover mapping: Preliminary results for northeastern Arizona. In: C. van Riper III and D.J. Mattson (Eds.). *The Colorado Plateau II: Biophysical, Socioeconomic, and Cultural Research*. The University of Arizona Press, Tucson, Arizona.

FAO. 2005. *Forest Resources Assessment.* FAO. Rome, Italy.

FOREST WATCH. 2010. http://www.globalforestwatch.org/ consulted on Jun 21, 2010.

FRIEDL, M. A., MCIVER, D. K., HODGES, J. C. F., ZHANG, X. Y., MUCHONEY, D., STRAHLER, A. H., WOODCOCK, C. E., GOPAL, S., SCHNEIDER, A., COOPER, A., BACCINI, A., GAO, F. & SCHAAF, C. 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment,* 83: 287-302.

GOPAL, S. & WOODCOCK, C, 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing,* 60(2): 181-188.

GUTIERREZ, L. A. 1992. *Diagnostico de las cuencas hidrográficas de Honduras.* Informe de Consultoria: Banco Interamericano de Desarrollo. Tegucigalpa, Honduras. 124 pp.

HARGREAVES, G. 1992. *Hydrometereologic data for Honduran water resources development.* USU, Dept. of Biological and Irrigation Engineering. Logan, Utah, 77 pp.

JENSEN, J. R. 2005. *Introductory digital image processing: a remote sensing perspective*. 3rd. Ed. Pearson-Prentice Hall. 526 pp.

LATIFOVIC, R., ZHU, Z., CIHLAR, J., GIRI, I. & OLTHOF, I. 2000. Land cover mapping of North and Central America-Global Land Cover. *Remote Sensing of Environment,* 89(2004): 116-127.

LABA, M. L., GREGORY, S., BRADEN, J., OGURCAK, D., HILL, R., FEGRAUS, E., FIORC, J. & DEGLORIA, S. D. 2002. Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sensing of Environment,* 82: 443-456.

LOTSCH, A., TIAN, Y., FRIEDL, M. A. & MYNENI, R. B. 2003. Land cover mapping in support of LAI and FPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. *International Journal of Remote Sensing*, 24(10): 1997-2016.

LOVELAND, T. R., REED, B. C., BROWN, J. F., OHLEN, D. O., ZHU, Z., YANG, L. & MERCHANT, J. W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *International Journal of Remote Sensing,* 21: 6, 1303-1330.

LOWRY, J., RAMSEY, R. D., THOMAS, K., SCHRUPP, D., SAJWAJ, R., KIRBY, J., WALLER, E., SCHRADER, S., FALZARANO, S., LANGS, L., MANIS, G., WALLACE, C., SCHULZ, K., COMER, P., POHS, K., REITH, W., VELASQUEZ, C., WOLK, B., KEPNER, W., BOYKIN, K., O'BRIEN, L., BRADFORD, D., THOMPSON, B. & PRIOR MAGEE, J. 2007. Mapping moderate-scale land-cover over very, large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP). *Remote Sensing of Environment,* 108: 59-73.

MUCHONEY, D., BORAK, J., CHI, H., FRIEDL, M., GOPAL, S., HODGES, J., MORROW, N. & STRAHLER, A. 2000. Application of the MODIS global supervised classification model to vegetation and land cover mapping of Central America. *International Journal of Remote Sensing*, 21(6): 1115 -1138.

ORDÓNEZ, T. M. M. & HOUSE, P. 2002. *Mapa de Ecosistemas Vegetales de Honduras: Manual de Consulta.* Proyecto P.A.A.R., Tegucigalpa, Honduras.

POIANI, K. A., RICHTER, B. D., ANDERSON, M. G., & RICHTER, H. E. 2000. Biodiversity Conservation at Multiple Scales: Functional Sites, Landscapes, and Networks. *Bioscience,* 50(2): 133-146.

RICHARDS, M. 1996. Protected areas, people and incentives in the search for sustainable forest conservation in Honduras. *Environmental Conservation* 23 (3): 207-217.

RULEQUEST RESEARCH. 2004. WWW URL: http://www.rulequest.com/index.html (accessed on June 21, 2010).

SHEPPARD, S. R. J. & CIZEK, P. 2009. The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualization. *Journal of Environmental Management*, 90(2009): 2102-2117.

STONICH, S. C. 1993. *I am destroying the land. The political ecology of poverty and environmental destruction in Honduras.* Westview Press, Boulder, Colorado, 204 pp.

UNESCO. 1973. *International Classification and Mapping of Vegetation.* United Nations Educational, Scientific and Cultural Organization, Paris, France.

VERMOTE, E. F. 2008. *MODIS Surface reflectance user's guide.* MODIS land use reflectance science computing facility. Prepared by E. F. Vermote, S. Y. Kotchenova and J. P. Ray. Web site: http://modis-sr.ltdri.org. 26 pp.