GIS and Remote Sensing Analyses of Forest Plantations in Costa Rica's Atlantic Lowland Area

Investigator: James Donahey[1]
Advisors: Marcia Snyder[2], Ann Russell[1], Lisa Schulte[1]
Affiliations: [1]Iowa State University, Department of Natural Resource Ecology and Management, Ames, IA USA
[2]Organization for Tropical Studies, La Selva Biological Station, Costa Rica

Abstract

With the recent implementation of the Kyoto Protocol (Kyoto 1997), carbon credits may afford a new economic market associated with forested landscapes (IPCC 2000). To assess the potential for carbon sequestration in an important part of those forested landscapes, tree plantations, it is not only important to determine the extent to which tree species differ in carbon storage, but also to know the areal extent of these plantations.

Here, two different types of images were used to develop a method for classifying and geo-referencing tree plantations in the Costa Rican lowlands. The first, an ALI multispectral image, yielded some success with 59% of the expected plantations found to be correctly identified. The second type, Hydice images with much smaller areal coverage, was found to be almost useless due to sensor noise.

For the multispectral ALI images, applying a Minimum Noise Fraction, MNF, filter to decorrelate noise and determine the dimensionality of the image was followed by a supervised Maximum Likelihood classification to distinguish plantations from primary and secondary forests in the surrounding areas. For the hyperspectral Hydice images, the MNF filter was again used followed instead by a Spectral Angle Mapper classification. The Spectral Angle Mapper was chosen due to this algorithm's ability to classify even with large variances in lighting (Lumme 2004). This method was found to be the most successful method of distinguishing plantations of differing species from the surrounding vegetation.

No definitive information on the areal extent of plantations in the Costa Rican lowlands was found, but a map of plantations that were ground-truthed was successfully created for future researchers to use more recent and higher resolution image sets, such as the CARTA II mission, to attempt similar or new methods.

Introduction

Dense tropical forests once covered much of Costa Rica. In a period from the 1960's to the 1980's 83% of Costa Rican forests were felled (Hartshorn 1992). The main factors driving this deforestation were a demand for agricultural lands such as pastureland, fruit plantations, and cropland (Veldkamp 1992). Reconversion to forests is a recent trend as local farmers, sometimes with the help of non-government organizations, establish tree plantations on former agricultural lands.

In 1980 Costa Rica had 2,800 hectares of land dedicated to forest plantations. The extent of forest plantations increased to 40,000 hectares by 1990 (FAO 1993). As of the year 2000, Costa Rica had 178,000 hectares of land used as forest plantations (FAO 2000). The dramatic increase in area of tree plantations has the potential to bring very important changes to the region's natural resource base. The benefits of these tree plantations may include improved soil and water quality, increased timber production, and improved wildlife habitat.

With the recent implementation of the Kyoto Protocol, carbon credits may afford a new economic market associated with forested landscapes (IPCC 2000). To assess the potential for carbon sequestration in tree plantations, however, it is important to determine the extent to which tree species differ in carbon storage in biomass and soil over the range of environmental conditions encountered within the region. If species do differ, then what is the extent of areal coverage required to observe measurable differences among species in carbon sinks at the regional level? To address these questions, an essential component is the development of the capability to identify and geo-reference the forest plantation land-use category, using satellite imagery and image processing software. When linked with process-based modeling, this capability will enable us to quantify differences among tree species in their potential to influence regional carbon storage.

The primary objective of my research was to develop and test a methodology to distinguish tree plantations of different species using remotely sensed images. My secondary objective was to delineate the current spatial extent of tree plantations in the Atlantic slope lowlands of Costa Rica.

Study Area

The main study site is a series of plots located on the Huertos peninsula at the confluence of the Sarapiqui and Puerto Viejo rivers within the borders of the La Selva Biological Research Station, Puerto Viejo De Sarapiqui, Costa Rica Longitude: 84°00'12.922"W, Latitude 10°25-52.610"N(OTS). The area is a tropical wet forest receiving an average of 4 m of rain per year. In addition to the main study site of the Huertos plots, a trial method was tested on an image covering a subset of (from the cities of Puerto Viejo De Sarapiqí to Trinidad and Canta Gallo to Selva) the Atlantic lowlands region of Costa Rica (Figure 1). This subset of the Atlantic lowlands was chosen due to the availability of a relatively cloud free EO-1 ALI image of the area.

Data Collection

Images of the region have been collected from a variety of sources. The most recent of those is an image from January 2004 of the Earth Observing (EO-1) satellite (January 2004 image). The EO-1's Advanced Land Imager (ALI) sensor captures images on 9 different bands with a spatial resolution of 30 meters, and a panchromatic band with a spatial resolution of 10 meters (Table 1). The EO-1 is spectrally similar to the Landsat satellites but provides a more current view of the regional land cover that was available from existing Landsat-7 ETM+ (June 2001).

Hydice hyperspectral images are also available for the northwestern region of La Selva. These images, taken in March 1998, contain 210 bands that capture wavelengths from .4µm to 2.5µm, and have approximately 1.5m resolution (OTS). The large spectral coverage and fine spectral resolution of the Hydice images allows for more information to be gathered from individual pixels and more detail to be visualized and classified than the EO-1

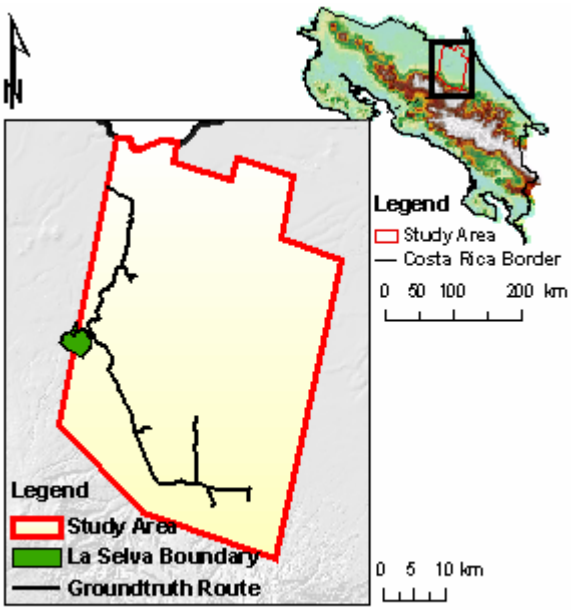

Figure 1.

The study area for the EO-1 image classification in relation to Costa Rica and La Selva.

ALI and Landsat-7 ETM+ images. Unfortunately, the plantations covered by the Hydice image are few and at the time of image capture were very young. As such, they do not accurately represent the plantation variety in the lowlands region and are of limited use in developing a methodology. The spatial extent of the Hydice images is much less than the EO-1 and Landsat-7 images, and has not been geo-referenced, but nonetheless the potential uses for such geo-referenced images with much larger coverage are tremendous.

Table 1. EO-1 ALI Capabilities (USGS)

Band	Wavelengh(µm)	Ground Sample Distance (m)
Pan	0.48 - 0.69	10
MS - 1	0.433 - 0.453	30
MS - 1	0.45 - 0.515	30
MS - 2	0.525 - 0.605	30
MS - 3	0.63 - 0.69	30
MS - 4	0.775 - 0.805	30
MS - 4'	0.845 - 0.89	30
MS - 5'	1.2 - 1.3	30
MS - 5	1.55 - 1.75	30
MS - 7	2.08 - 2.35	30

The EO-1 image was geo-referenced when it was captured, but the spatial offset of greater than 150 meters was too inaccurate for our purposes. A combination of thirty-one ground control points derived from a GPS with wide area augmentation system, WAAS, enabled and a Landsat-7 image were used to geo-reference the image. A 30 meter digital elevation model (DEM) from the Instituto Geographica National of Costa Rica was used to orthorectify the EO-1 ALI image. Due to the time constraints of the project, the Hydice images were classified in their non-geo-referenced form.

Methods

ALI image preparation

Before analysis could be attempted, some preprocessing work was performed on the images. Some of the ALI bands contain highly correlated, and therefore redundant, information. Using Principal Components Analysis (PCA; Keshava and Mustard 2002) the dimensionality in the image was decreased and the three most relevant bands remained. It was after this that the first signs of a signature difference between tree plantations and secondary forest, and the similarity between the spectral signatures of tree plantations and agricultural land, such as banana and pineapple plantations, was first observed. An example of the signature differences between forest and tree plantations can be seen in Figure 2. To augment the difference in spectral signatures more transformations were performed.

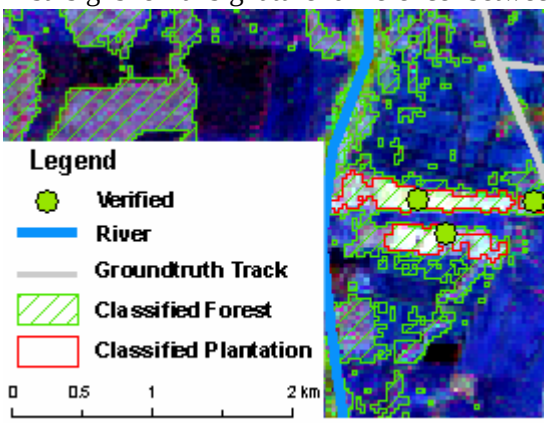

Figure 2.

A sample classification of the Atlantic lowlands. Green areas were classified as forest, and the area highlighted in red was classified as a plantation. The variation in signature between the two is visible.

ALI image classification

Due to the observed spectral similarities between tree plantations and certain agricultural lands, such as banana and pineapple plantations, a mask was created to separate the forested areas from the non-forested areas. Unsupervised ISODATA classifications were attempted on the original unprocessed ALI image with varying

ranges of classes. This always seemed to result in too few or too many classes.

The most accurate classification for distinguishing the classes desired was a supervised maximum likelihood classification. The training areas used were forested area, forest shadow, water, pasture, banana plantations, and known forest plantations. The maximum likelihood classification is used when the statistics for each class in each band are assumed to be normally distributed. With each iteration, each pixel is assigned a class to which it has the highest probability of belonging (Richards 1999). The non-forested area classifications were merged and the forested area, forest shadow, and tree plantation classes were also merged. A majority filter was used to remove some of the smaller patches of isolated trees from the image. In this majority filter process, a pixel is replaced with the value of the majority of the pixels in a surrounding grid. The grid size was set to 3 x 3 pixels. The end result was a mask of all of the forested areas in the image with the smaller, isolated, patches of trees removed.

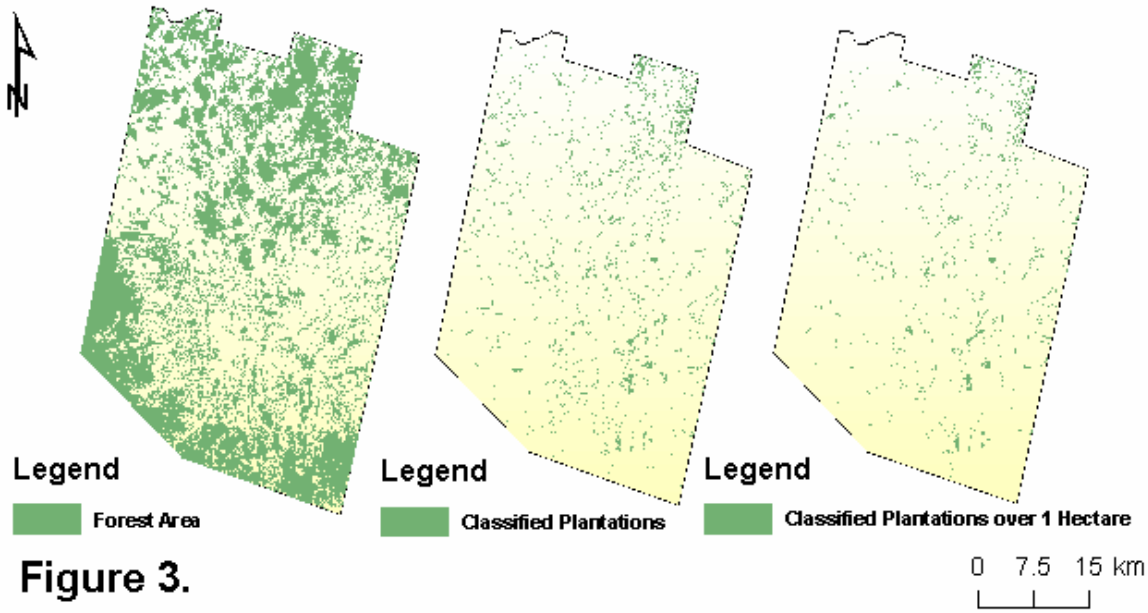

Figure 3.

Results of a series of classification of the Atlantic lowlands. Green areas in the first map were classified as forest and used as a mask to create the classification of plantations in the next image. The green areas highlighted as possible plantations were then filtered for size and all plantations greater than a hectare in size remained.

Using the Minimum Noise Fraction rotation (MNF; Green et al. 1998), the original EO-1 ALI image was transformed. This method performs two Principal Components Analysis transformations to decorrelate noise between the bands and determine the dimensionality of the image. After the MNF transformation, the number of bands in the image was left unchanged but those that remained contained decorrelated information. Eight of the new bands were relatively free of noise with a ninth band that was mostly noise. Two of the noise free bands appeared to show the most contrast between the tree plantations and forests.

After several attempts of maximum likelihood classifications using different combinations of transformed bands with varying contrasts between tree plantations and forests, a single band with the most contrast between the two classes was selected that made the most accurate classification. Training areas in known plantations and known forest were selected to run a supervised maximum likelihood classification. The training areas used were known tree plantations, forest, and forest shade. This classification was subset to the forest/non-forest vector image. After the classification, the classes of forest and forest shade were deleted, and the remaining area of plantations was left. Due to the similarity in signature between isolated trees

and plantations, a majority filter was used to remove the isolated trees that showed up as suspect plantations. Again, the grid area selected was 3 x 3 pixels. The area for the potential plantations was calculated and areas less than one hectare were filtered out. The final map of suspect plantations was created from this classification (Figure 3). Points were given to each suspect plantation, and the areas were ground-truthed using a Garmin GPS.

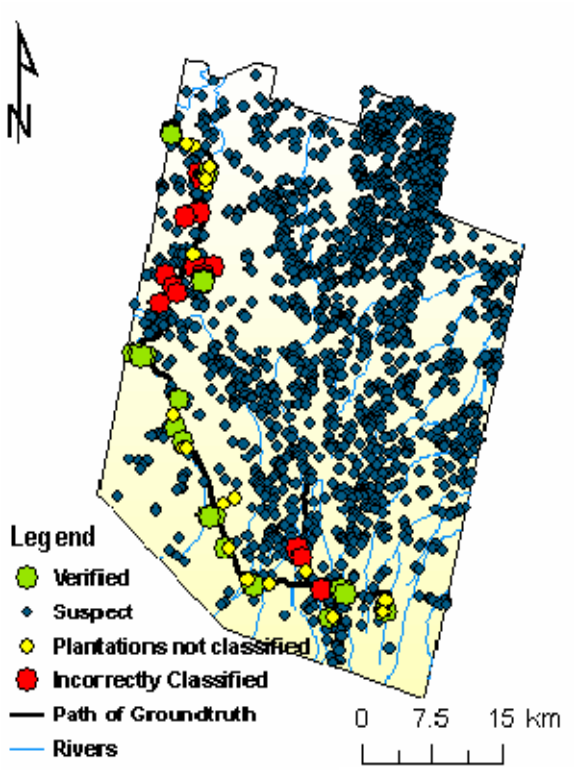

Figure 4.

Potential and verified plantations within the study area. The ground-truth track is also highlighted. A band is visible that may be related to sensor issues.

Ground-truthing

With the advice of other researchers in the region, we drove to known forest plantations and marked their locations using a Garmin GPS unit with WAAS enabled. Along the same roads we searched for plantations that were categorized from the classification (Figure 4). In addition to the locations of the new plantations, we recorded the species, if the plantation appeared to be <5 yrs old, and the direction from the road that the plantation was located.

HYDICE image preparation

To compare classification accuracy to hyperspectral images with higher resolution I tested a similar method on the available Hydice images covering the Huertos plots within the boundary of La Selva (Figure 5). Unfortunately, the Hydice images contain a considerable amount of noise due to some unknown sensor problems. The maximum noise fraction transform was again used to separate the bands containing noise and those containing relevant information. Instead of keeping all of the bands, as in the ALI image, I chose to keep 45 of the 210 bands. These bands explained about 66% of the variance in the pixel statistics. Even then, only 12 of the bands were relatively noise free.

HYDICE image classification

In place of the maximum likelihood supervised classification, this time a spectral angle mapper classification was run. One of the difficulties faced with the maximum likelihood classifier is that each category in the training areas must contain at least as many pixels as bands. These training areas should be homogenous areas that are representative of the category that is attempting to be classified. With the many bands in the hyperspectral images, it was difficult to find smooth homogenous areas to fulfill the requirements of a maximum likelihood classification.

The spectral angle mapper algorithm does not have a minimum number of pixels for each class to run a successful classification, and also works better with variances in lighting (Lumme 2004). Due to the large amounts of shadows, and small areas of plantations, this algorithm had the potential to create a more accurate classification. The spectral angle mapper algorithm compares the angle between the endmember spectrum and each pixel vector in n-

dimensional space. Instead of comparing the intensity of reflection of each pixel, it compares the angle at which the vector is plotted. Smaller angles represent closer matches to the reference spectrum. (Kruse et al. 1993)

The results from the classification were filtered using a majority filter of a 3 x 3 grid to remove smaller patches, and a clumping filter was used, also with a grid of 3 x 3 pixels, to make the obvious groups more solid (Figure 6). Ground-truthing was not necessary due to the known established boundaries of the Huertos plantations (Figure 5).

Results

ALI Classification

Of the 1600 suspect plantations in the study area, 34 areas were ground-truthed on three separate days. Twenty areas, or 59% of the ground-truthed areas, were correctly classified as plantations from the ALI image. These areas varied in age and included 8 of the species found to be planted in the area. Fourteen areas, or 41% of the ground-truthed areas, were incorrectly classified as plantations. These areas were found to be disturbed primary forest in 4 locations and secondary forest at 10 locations. Due to restrictions on time, the poor quality of roads, and the large study area, more ground-truthing did not occur.

Nineteen other plantations were found while ground-truthing. These locations were also marked. Six of these areas were found after the fact to have been accurately classified as plantations before the classified areas were filtered for everything smaller than a hectare in size. The plantations that were not found until ground-truthing were of similar species to those that were found with the classification. The smallest plantation classified correctly as a plantation was 11700 m^2, and the smallest classified area in plantation was 900 m^2.

Hydice Classification

Hydice images were not geo-referenced and were not ground-truthed. Using established maps of the Huertos plots, classified areas were compared to the known boundaries. Without geo-referencing, filtering for size was not possible, and without being able to complete the procedure a percentage error is not known for these areas.

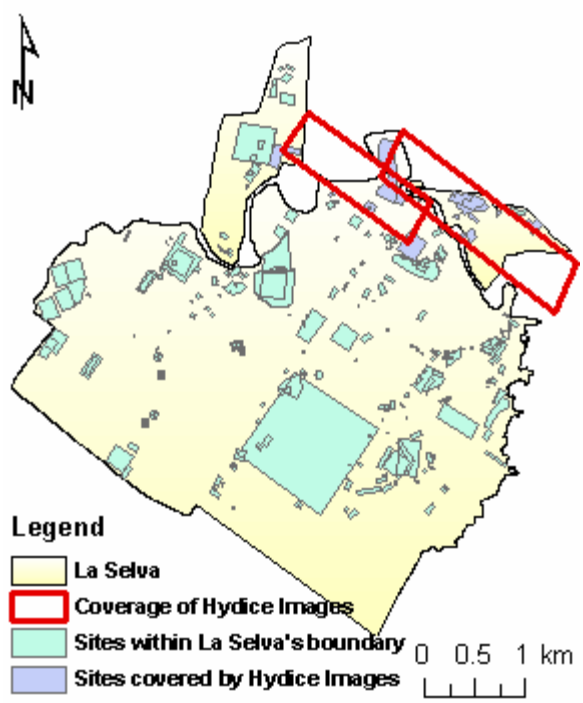

Figure 5.

Project areas within the boundaries of La Selva. The sites highlighted in violet are those with some portion being covered by Hydice images.

Discussion

ALI Classification

After the first classification using the MNF transform, it was obvious that the classification was not only picking up tree plantations, but smaller, isolated, forests as well. Some of these areas were riparian areas, and others were isolated trees. There could be several reasons for the spectral similarities between these areas, tree plantations, and fruit plantations. The growth rates of isolated trees, riparian trees,

and tree plantations, may be more than those of trees facing greater competition in secondary and primary forests. This may be creating a unique spectral signature that is separable after running the MNF transformations, and shared by these types of vegetation. In 14 cases, secondary forest and disturbed forest were also classified as plantations. Some of these areas, if disturbed or regenerated in a certain way, might also reflect this similar signature.

In order to correct these problems using the same image set, the possibility of clipping plantations to a specified buffer from roadways might be a practical way of reducing incorrect classifications. Because the plantations are human made, their locations tend to be near roads. However, there is no practical way of removing buffering around rivers. Several plantations were found to be planted directly bordering a river.

Using the relatively arbitrary size of one hectare did remove six of the plantations found with the help of other researchers on our ground-truthing expeditions that would have been successfully classified using the above methods. This filter also removed many smaller isolated groups and narrowed down our search for plantations. Many of the plantations that we visited were <1 hectare, but with the pixel size of the EO-1 ALI images and the spectral mixing that occurs, it is very difficult to accurately classify small areas.

The 30-meter resolution was the primary limiting factor with the EO-1 ALI satellite images. The spectral mixing of areas made separating the relatively small plantations very difficult and made the already few bands even more useless. When the spectral preprocessing techniques were performed on the ALI image, it was also noticed that there were problems with at least one of the sensor bands in the satellite. This might have led to the appearance of a well-defined streak of suspect plantations in the final classifications.

After three days of ground-truthing, the tree plantation points collected were used to locate the plantations on the EO-1 ALI image and create a polygon for each tree plantation. These ground-truthed plantations will serve as a larger database of tree plantations to use for conducting the classifications at a later date, with a better image set, and to aid researchers in the area who are looking for plantations that they might use as future study cites.

Hydice Classification

Even with the additional bands of information and the higher resolution the Hydice images, classifying these images did not yield any better results. The large amount of noise in each image made them almost useless. Hyperspectral noise tended to be more correlated than actual objects in the images. Maximum Noise Fraction transforms explained 95% of the variation between pixel values resulted in 185 bands, only a dozen of which were noise-free. Even though continuous shapes are visible in the areas of plantations, they are also visible in areas where there were no plantations at all. Shadows, clumps of similar tree species, and noise were the most likely cause of these areas. In areas of plantations there is so much variation in signals, due to the high spatial resolution of the imager, that it is difficult to get continuous areas. More spectral mixing, with a lower resolution imager, would be desired.

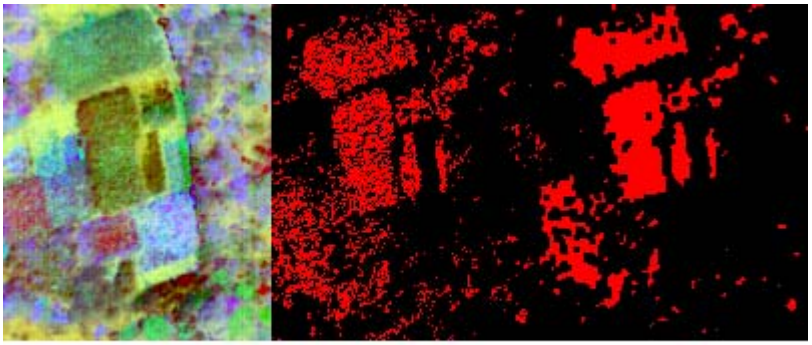

Figure 6. Classified Plantation

A series of the Hydice classifications beginning with the original image. The second image is the first results of a classification, and the final image is after the majority filter has been performed

As with the ALI image, unsupervised ISODATA classifications were attempted, but again

never resulted in the desired output. There always appeared to be either too many categories or too few categories. If too many categories were used than there were no clear delineations between plantations. If too few, then the plantations simply faded into the surrounding secondary forest.

Because of the smaller coverage area, a forest/non-forest map wasn't necessary to separate between a tree plantation and a banana/pineapple plantation. To test the possibility of running such a classification, however, a two band buffer around red, green, blue, and infrared regions on the electromagnetic spectrum was used for a six category supervised maximum likelihood classification. The classes were: water, grass; sand; forest; forest shade; and tree plantations. All of the non-forested area classes were merged, and a successful mask of forest and non-forest areas was created.

If the images were geo-referenced, the resulting areas of suspect plantations can be filtered for size and ground-truthed as for the ALI image. The only plantations visible in the images were the Huertos plots, and some outside of La Selva's boundary. These plantations were relatively young when the pictures were taken, in 1998, and there is not as much diversity among the plantation types that you would see on an image of a larger spatial extent.

General Discussion

The main obstacle that future researchers might have to address is one that technology can not necessarily answer. In order to gain an accurate view of plantations, careful ground-truthing with the permission of the landowner would be necessary. For any future missions to be successful they must form agreements with the landowners and local NGO's that have led the way in encouraging tree plantations as an alternative to local farmers. Tree plantations in the region are not always one large monoculture plot. Many of the plantations that were ground-truthed were found to be many small plots of varying species planted together to form one large plantation. Combined with the low spectral and spatial resolution of the EO-1 ALI image, this made identifying plantations even more difficult. In order to establish definite boundaries for species, if species level identification is desired, these agreements with landowners must be made.

Two methods for classifying plantations by species should be explored with future image sets. The first is to use the hyperspectral images of a high spatial resolution, somewhere between 2-10 meters, to run a classification with different training classes in different known species plantations. After running a classification the results can be clumped together to form definite shapes and the areas can be filtered for size to find areas of plantations.

The second, which was investigated using the ALI images, is to run a classification using training areas in plantations of all species and then reclassify those areas by species at a later time with images of both a greater spatial and spectral resolution. Either method has its shortcomings, and current image coverages do not exist which would allow both to be done accurately. Most multispectral images have larger coverage, but lack the spatial and spectral detail to classify plantations by species. Hyperspectral images may have finer spatial and spectral resolution, in this case, but almost have too much information to pull out clumps like plantations. With OTS's recent purchase of the Carta II mission, which covers almost the entire country, future more accurate classifications would be more possible if one could address the problems associated with high-resolution imagery.

References

FAO. 1993. Forest resources assessment 1990. Tropical countries. FAO Forestry Paper 112. Rome, 59 pp.

FAO. 2000. Global Forest Resources Assessment 2000. Central America and Mexico

FAO Forestry Paper - 140 <http://www.fao.org/documents/show_cdr.asp?url_file=/DOCREP/004/Y1997E/y1997e14.htm>.

IPCC (Intergovernmental Panel on Climate Change). 2000. IPCC Special Report on Land Use, Land-Use Change And Forestry. < http://www.ipcc.ch/pub/reports.htm>

Keshava, Nirmal., and Mustard, John., "Spectral Unmixing", IEEE Signal Processing, Exploiting Hyperspectral Imagery Vol 19, pg 44-57, Jan 2002.

Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P. & Goetz, A. 1993. The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2-3), pp. 145-163.

Kyoto 1997, Protocol to the United Nations Framework Convention on Climage Change United Nations Framework Convention on Climate Change <http://unfccc.int/essential_background/kyoto_protocol/background/items/1351.php>

Lumme, J. H. , 2004, "Classification of vegetation and soil using imaging spectrometer data". Proc. XXth ISPRS Congress, 12-23 July 2004 (Istanbul, Turkey)

OTS. La Selva GIS, <http://sloth.ots.ac.cr/~msnyder/laselva_gis/index.html>. 28 Dec. 2005

OTS. La Selva Images, <http://www.ots.ac.cr/en/laselva/gis/laselva_gis/frames/lsimages.htm>. 22 Oct. 2005

Richards, J.A., 1999, *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlin, p. 240.