The present invention provides systems and methods to automatically analyze Landsat satellite data of forests. The present invention can easily be used to monitor any type of forest disturbance such as from selective logging, agriculture, cattle ranching, natural hazards (fire, wind events, storms), etc. The present invention provides a large-scale, high-resolution, automated remote sensing analysis of such disturbances.

17 Claims, 10 Drawing Sheets
FIG. 5
FIG. 8
For the purpose of illustrating the invention, there are depicted in the drawings certain embodiments of the invention. However, the invention is not limited to the precise arrangements and instrumentalities of the embodiments depicted in the drawings.

FIG. 1 depicts spatial distribution of selective logging in five timber production states of the Brazilian Amazon for the year intervals 1999-2000 (red), 2000-2001 (blue), and 2001-2002 (green). The states of Amazonas (AM), Amapa (AP), Tocantins (TO), Maranhao (MA), and the southern non-forested part of Mato Grosso were not included in the analysis. Light gray areas show the extent of Indigenous reserves; dark gray areas delineate federal conservation lands as of 1999 Instituto-Socioambiental. (São Paulo, Brazil, 1999) Map of forest types, land-use change and protected areas in the Amazon.

FIG. 2 depicts a high resolution example of selective logging results in 2001-2002 from the CLAS processing in comparison to deforestation mapping provided by the Brazilian National Institute for Space Research (INPE) (Instituto Nacional de Pesquisas Espaciais). "PRODES: Assessment of Deforestation in Brazilian Amazonia (http://www.obt.inpe.br/prodes/index.html)" (2005)).

FIG. 3 depicts the Carnegie Landsat Analysis System (CLAS) processing stream.

FIG. 4 depicts the AutoMCU sub-model within CLAS, showing that each satellite image pixel is a calibrated reflectance spectrum that is deconvolved into constituent fractional covers of photosynthetic vegetation (PV), non-photosynthetic vegetation (NPV), and soil. Spectral endmember libraries developed from extensive field and hyperspectral satellite studies (TropiSpec) (Asner et al., 2005) are in a probabilistic Monte Carlo unmixing approach to derive the percentage cover of PV, NPV, and soil within each image pixel.

FIG. 5 depicts spectral endmember bundles used in the AutoMCU step of CLAS (from FIG. 3), which are (A) Photosynthetic vegetation, (B) Non-photosynthetic vegetation, and (C) soil. Adapted from Asner et al. (2004a).

FIG. 6 depicts an example of deforestation and water body masking using Landsat thermal band 6 and the AutoMCU result for photosynthetic vegetation (PV).

FIG. 7 depicts an example of logging detection using CLAS. AutoMCU results from one year are differentiated against those of the next year. A directional pattern recognition algorithm then uses the PV-change image to locate probable logging decks, skids, and roads.

FIG. 8 depicts a geographic coverage of study, showing the Brazilian Legal Amazon with Landsat 7 satellite footprints.

FIG. 9 depicts an example showing how the CLAS logging product is unique from the PRODES deforestation products provided by the Brazilian Space Research Institute.

FIG. 10 depicts a block diagram of the CLAS system.

DETAILED DESCRIPTION OF THE INVENTION

All publications and patent applications herein are incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

The following description includes information that may be useful in understanding the present invention. It is not an admission that any of the information provided herein is prior art or relevant to the presently claimed inventions, or that any publication specifically or implicitly referenced is prior art.
Selective logging rates from 1999-2002 in five major timber-producing states of the Brazilian Amazon, with comparison to the deforestation rates reported by INPE (2005).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(km² yr⁻¹)</td>
<td>(km² yr⁻¹)</td>
<td>(km² yr⁻¹)</td>
</tr>
<tr>
<td>Acre</td>
<td>64</td>
<td>547</td>
<td>53</td>
</tr>
<tr>
<td>Mato Grosso</td>
<td>13,843</td>
<td>6,176</td>
<td>7,912</td>
</tr>
<tr>
<td>Rondonia</td>
<td>5,939</td>
<td>6,671</td>
<td>5,234</td>
</tr>
<tr>
<td>Pará</td>
<td>773</td>
<td>2,465</td>
<td>3,237</td>
</tr>
<tr>
<td>Roraima</td>
<td>52</td>
<td>253</td>
<td>55</td>
</tr>
<tr>
<td>Total</td>
<td>20,651</td>
<td>16,112</td>
<td>14,286</td>
</tr>
</tbody>
</table>

*Only the northern 58% of Mato Grosso containing forested lands was included in the analysis.

Selective logging doubles previous estimates of the total amount of forest degraded by human activities (Table 1), a result with potentially far-reaching implications for the ecology of the Amazon forest and the sustainability of the human enterprise in the region. In the future, improved monitoring of tropical forests will require high performance satellite observations and new computational techniques. The results, presented with explicit uncertainty analysis and transparency of
method, have located and quantified ubiquitous but previously cryptic disturbances caused by selective logging. Definitions.

As used herein, each of the following terms has the meaning associated with it in this section.

The articles “a” and “an” are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

As used herein, the term “deforestation” refers to clearing and conversion of the forest to other land uses, such as cattle pasture, crop agriculture, and urban and suburban areas.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXPERIMENTAL EXAMPLES

The invention is now described with reference to the following examples. These examples are provided for the purpose of illustration only and the invention should in no way be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

The materials and methods used in the experiments presented in this Example are now described.

Materials and Methods

Processing Methodology

The Carnegie Landsat Analysis System (CLAS) includes a general purpose computer programmed to use high spatial resolution satellite data for regional and global studies of forest disturbance. The computer system used is a multi-processor Linux system, but other systems can be used. CLAS is an automated processing system that includes: (i) atmospheric correction of satellite data; (ii) deconvolution of spectral signatures into sub-pixel fractional cover of live forest canopy, forest debris and bare substrates; (iii) cloud, water, and deforestation masking; and (iv) pattern recognition algorithms for forest disturbance mapping. The following sections provide a detailed description of CLAS, illustrated by Fig. 3.

Image Preparation and Atmospheric Correction

The version of CLAS presented herein ingests raw Landsat Enhanced Thematic Mapper Plus (ETM+) satellite imagery and applies sensor gains and offsets to convert from digital number (DN) to exo-atmospheric radiance. The radiance data are passed to a fully automated version of the 6S atmospheric radiative transfer model (Vermote et al.). The 6S program is integrated into the CLAS processing stream and uses monthly averages of aerosol optical thickness (AOT) and water vapor (WV) values from the Moderate Resolution Imaging Spectrometer (MODIS) sensor onboard the NASA Terra spacecraft. Time-stamping of MODIS AOT and WV data with Landsat data is done on an automated basis (Fig. 3).

Sub-Pixel Analysis

The CLAS process relies upon the quantitative determination of fractional material cover at the sub-pixel scale (e.g., within each Landsat 30×30 m pixel). This core step employs a probabilistic spectral mixture sub-model that is run using the formulation shown in Fig. 4. This process spectrally decomposes each image pixel into fractional cover estimates (0-100% cover) of photosynthetic vegetation (PV) canopy, non-photosynthetic vegetation (NPV), and bare substrate. This sub-model is based on an algorithm developed for forest, savanna, woodland and shrubland ecosystems. It is fully automated and uses a Monte Carlo Unmixing (AutoMCU) approach to derive uncertainty estimates of the sub-pixel cover fraction values. The method uses three spectral end-member “bundles”, derived from extensive field databases and satellite imagery, to decompose each image pixel using the following linear equation:

\[
\rho(\lambda)_{pixel} = \sum(C_p, \rho(\lambda)_p) + C_{non-photosynth} \rho(\lambda)_{non-photosynth} + C_{bare} \rho(\lambda)_{bare} + \epsilon
\]

(i)

where \(\rho(\lambda)_p\) is the reflectance of each land-cover endmember (e) at wavelength \(\lambda\) and \(\epsilon\) is an error term. Solving for the sub-pixel cover fractions \(C_p\) requires that the observations \((\rho(\lambda)_{pixel})\) in this case, Landsat ETM+ reflectance) contain sufficient spectral information to solve a set of linear equations, each of the form in equation (1) but at different wavelengths (\(\lambda\)).

Until recently, there were a limited number of spectral signatures of green and senescent vegetation and bare substrates for tropical regions. The mixture modeling technique requires spectral reflectance bundles \(\rho(\lambda)_p\) that encompass the common variation in canopy and soil properties. Asner (1998) and Asner et al. (2003a, 2004a) collected these spectral data using full optical range field spectroradiometers (Analytical Spectral Devices, Inc., Boulder, Colo., USA) during field campaigns conducted from 1996 to 2000. The spectral endmember database encompasses the common variation in materials found throughout the Brazilian Amazon, with statistical variability well defined (2004a). The bare substrate spectra have been collected across a diverse range of soil types, surface organic matter levels, and moisture conditions. Spectral collections for NPV have included surface litter, senescent grasslands, and deforestation residues (slash) from a wide range of species and decomposition stages.

In contrast to the NPV and bare substrate spectra that can be collected via ground-based spectroscopic measurements, the photosynthetic vegetation (PV) spectra of forest species require overhead viewing conditions. This is very difficult in forest canopies with heights typically ranging from 10-50 m. Spectral measurements of individual leaves, stacks of foliage, or partial canopies (e.g., branches) introduce major errors in spectral mixture models and cannot be used (Asner, 1998). Therefore, canopy spectra were collected using the Earth Observing-1 (EO-1) Hyperion sensor, the first spaceborne hyperspectral sensor for environmental applications (Ungar et al.). The PV spectral bundle was derived from more than 40,000 spectral observations made at 30 m spatial resolution with Hyperion (images taken throughout 1999), atmospherically corrected to apparent top-of-canopy reflectance using the ACORN-4 atmospheric correction algorithm for hyperspectral data (ImSpec Inc., Palmdale, Calif. USA), and convolved to Landsat ETM+ optical channels (Asner et al., 2005). These green vegetation spectra thus inherently included the variable effects of intra- and inter-crown shading, which are prevalent in tropical forests (Gastellu-Etchegorry et al.). In Amazonia, shade fractions average 25% cover in humid tropical forests, but the variance is high with standard deviations of 12% or more (Asner et al., 2003b).

It is thus critically important to note that the PV results include shade, which varies substantially with forest structure. Using a separate shade endmember is attractive (Souza et al., 2000), but doing so with multi-spectral Landsat data and such high shadow fraction variability often results in an under-determined spectral and mathematical problem in lin-
ear mixture models. That is, there are many viable solutions to the mixture modeling problem in forests. Imaging spectroscopy (hyperspectral) data are needed to solve this problem (Robert et al., 1993). This issue was avoided by accepting the limitations of incorporating variable shade directly into the PV bundle derived from the EO-1 Hyperion sampling of undisturbed forest canopies in Brazil. The PV bundle includes spectra from mature forest, late-stage forest regrowth, and logged forest of at least five years post-harvest. In the end, the total number of spectra retained in the end-member bundles for the AutoMCU sub-model was 252, 611, and 434 for PV, NPV, and bare substrate, respectively (FIG. 5). These spectra represent more than 130,000 field and spaceborne spectrometer observations collected over a five-year period of study (Asner et al., 2005).

Non-Forest Masking and Atmospheric Compensation

A thorough atmospheric correction was designed to exclude clouds, water bodies, cloud shadows, non-image and non-forest areas (e.g., pasture, urban and agriculture) from the CLAS processing stream (FIG. 3). Prior to execution of the AutoMCU sub-model, clouds are masked using the thermal channel (band 6) from the raw Landsat images. Asner et al. (2005) found that a thermal band threshold DN value of 125 can conservatively detect cloudy pixels over Amazonia. Water bodies are masked by finding pixels in the calibrated Landsat reflectance data in which bands 1-4 (blue, green, red, and near-infrared) have a negative slope. Only water displays such a negative reflectance slope with increasing wavelength. Non-image areas containing zero values are also masked.

Cloud shadows are identified using the root mean square error (RMSE) image that results from the AutoMCU processing (FIGS. 3-4). Areas shadowed by clouds have large RMSE values and are masked by identifying pixels above a specific RMSE threshold (Asner et al., 2005). To limit the logging analysis to forested areas, Landsat thermal band 6, combined with the AutoMCU results, is used to identify points containing primarily forest and non-forest areas. Forests have a lower brightness temperature and a higher PV fractional cover than deforested lands. A conservative PV fractional cover threshold of 60% was employed to delineate forest cover in the PV mask. The minimum and maximum thermal thresholds, which encompass forested areas in the thermal mask, are dynamically generated for each forest by calculating the mean thermal value of all pixels having a PV fraction cover greater than 80% and then masking all pixels with values >15 digital numbers (DN) from the mean thermal value. These final masking steps have the added feature of removing residual clouds and cloud shadows that were masked earlier in the CLAS processing (FIG. 6).

Although atmospheric correction was performed on the raw imagery before processing through the AutoMCU sub-model, residual atmospheric effects can persist (Asner et al., 2005). These residual effects exist spatially within a scene and temporally between scenes. These effects were greatly reduced prior to automated logging detection (next section) by calculating the average change in fractional forest cover in 55 km² subsets of the imagery. These large geographic subsets are made at a spatial scale far greater than that of the most extensive logging activities, so temporal differences in the overall forest fractional cover at this scale are a result of atmospheric effects (e.g., haze) or forest phenology. These false fractional cover changes are normalized by adjusting the background forest temporal variation to zero. Since disturbances related to logging or other anthropogenic activities occur at a much smaller spatial scale than is considered in this processing step, normalization of the forest values across large areas does not affect the CLAS process in discriminating true disturbances from the surrounding forested areas. Pattern Recognition

The specific criteria used in this procedure were determined following a comprehensive analysis and review of the forest responses to logging at various intensities in the Brazilian states of Pará, Mato Grosso and Acre where field studies were conducted. The mean and standard deviation fractional cover images from the AutoMCU step in CLAS provide quantitative data on canopy damage and forest disturbance intensity from which selectively logged areas can be determined (FIG. 3). By identifying areas of canopy disturbance that are arranged in specific spatial patterns, it is possible to detect logged areas on an automated basis. The primary method by which logging is detected is image differencing, where pairs of AutoMCU sub-pixel fractional cover images, separated by approximately one year, are used to create images of PV (forest canopy) and NPV (surface woody and senescent vegetation material) change that indicate areas of relative canopy disturbance or recovery. Forest disturbances in these images always have reductions in PV, simultaneous with increases in NPV fractional cover.

Logging activity results in low intensity forest disturbances from tree felling gaps, moderate intensity linear features from skid trails along which felled trees are dragged by tractors or skidders, and high intensity points of damage called log decks where logs are loaded onto trucks for transportation. The log decks are connected by logging roads, seen as linear features causing large reductions in the fractional cover of PV, to local roads or rivers for transportation to markets. These patterns are unique to logging throughout most of the Amazon, and thus they serve as the basis upon which the method for logging detection functions. CLAS identifies points (e.g., treefall gaps and log decks) and linear features (e.g., skid trails and logging roads) of recent disturbance occurring in forested areas. As these features also exist at a lower frequency in intact forest regions, their spatial density and diversity (see definition in next section) are calculated to identify those areas having disturbances in patterns most indicative of logging activity. The procedure then identifies these areas for further analysis by creating point maps, termed logging nodes, indicating their locations.

Log decks are automatically detected by searching for pixels where PV decreases significantly in a 30 m pixel centered on a 7×7 pixel kernel (4.41 ha). A positive detection is flagged when pixels with large PV reduction are surrounded by three concentric rings of incrementally greater PV cover surrounding the target pixel. This indicates an increase in canopy damage with greater proximity to the log deck, a pattern consistent with most logging activities.

The strategy for detecting decks works well in areas logged at higher intensities, as the decks tend to be abundant and equally spaced. However, in areas where the logging is more haphazard, where the forest damage is extremely high or low, or where the roads themselves also function as loading zones, individual log decks are not always distinguishable. Skids trails are a typological feature of selective logging practices, and they are the single-most ubiquitous surface feature found in harvested areas (Pereira et al., 2002; Asner et al., 2004).

The presence of skid trails is quantifiable based on large decreases in PV fractional cover in linear or near-linear patterns (Asner et al., 2004a). To detect the concentration of skid trails and auxiliary roads, a moving 6×6 pixel (3.24 ha) kernel is applied to the PV change image to enhance linear features in the N-S, E-W, NE-SW, and NW-SE directions (FIG. 3). The number of directions in which the linear features are arranged (which are defined herein as their diversity), and
their spatial density, in conjunction with the presence or absence of logging decks, is calculated for each location. With this information, it is possible to automatically distinguish probable logging events. In general, areas of greater logging intensity have a roughly equal proportion and higher density of linear features with the presence of logging decks. Lower intensity areas are normally dominated by one direction of linear feature and have few or no logging decks. An example of a typical logging detection is shown in FIG. 7.

Final Integration

After the linear and logging deck pattern recognition steps are completed, CLAS automatically integrates the various results to identify contiguous pixel clusters of probable logging activity. This process starts by creating a list of the logging nodes that are identified in the previous steps. Logged areas are identified using a moving kernel approach. A base kernel of 7x7 pixels (4.41 ha) and four 3x3 pixels (0.81 ha) subset kernels, one located at each corner of the base kernel, are used. The base kernel begins at each logging node and tests the criteria described below. If the area in question tests positive, the analysis kernel is moved to its 7x7 pixel neighbors to the north, south, east, and west, which are then each tested against the criteria (FIG. 3). This iterative process continues until all neighbors have been evaluated or the maximum logged cluster size (maximum of 17 positive detections per logging node) has been reached. The input layers and specific criteria tested within the base and subset kernels are described below. For the criteria below, all units for PV and NPV are % fractional cover within a pixel; units for PV CI and NPV CI are % change in cover fractions between image dates.

Input Layers to Logged Area Detection Procedure:

Logging node map
Thermal RMS mask (dynamically generated in earlier procedure) (T-mask)
PV mask (>60% fractional cover) (PV-mask)
NPV change difference image (NPV CI)
NPV change difference image (NPV CI)
After image PV (AI PV)

Base Kernel Criteria:
75% good data pixels (not cloud, cloud shadow, or water)
Non-forested area <0.54 ha (12.2%); based on T- and PV-masks.
60%<Mean AI PV>93%
Mean PV CI>9%
Mean NPV CI>2%
Mean NPV CI standard deviation>33%
Mean NPV CI standard deviation>46%
More than 6 pixels (0.54 ha) with PV CI values>80%
More than 6 pixels (0.54 ha) with NPV CI values<-85%
Masked area<0.18 ha

Subset Kernel Criteria:
±2 subsets with PV CI≥32% standard deviation
±2 subsets with mean PV CI≥3% and ±60%
±2 subsets having ±1 pixel (0.09 ha) with a PV CI value≥80%
±2 subsets with NPV CI≥46% standard deviation
±2 subsets with mean NPV CI≥5% and ±65%
±2 subsets having ±1 pixel (0.09 ha) with a NPV CI value≥85%

Manual Audit
Maps of probable logging events were visually audited to verify whether an area is being logged or not, in accordance with criteria established for identification of logged areas (see criteria below). In this process, false positives and negatives were manually removed and added. In this Amazon study, two analysts were employed during the audit, and their results and uncertainties were monitored and compared.
must be standardized. The annualization methodology used by the Monitoring the Brazilian Amazon Gross Deforestation Project (PRODES) was adopted. Assuming that logging, like deforestation, occurs during the dry season, the annual logging rate was determined by prorating the amount of logged area to a complete dry season. This standardization depends upon the onset and length of the dry season for each image, based on the geographic location of the scene center. The general spatial pattern of the Amazon dry season was provided by Marengo et al. (2001), based on the daily amount of outgoing longwave radiation.

Geographic Coverage

The Brazilian Amazon basin covers an area of approximately 4.1 million km². Analysis of the entire region with Landsat Enhanced Thematic Mapper-Plus (ETM+) imagery would require approximately 220 scenes per year or 880 images for the years 1999-2002, yet much of the northwestern Amazon still contains relatively little deforestation and logging (Nepstad et al., 1999). Therefore, the study was limited to the States of Acre, Pará, Mato Grosso (northern 58% of the state containing most of the forested area), Rondônia, and Roraima (FIG. 8). These five states contain ~90% of the deforestation reported by Brazil for all of the Legal Amazon, and thus are the most important areas for logging studies today. This strategy reduced the number of required Landsat images to 480 scenes.

The geographic overlap was evaluated between the Brazilian Amazon basin and the CLAS logging maps for all Landsat images used in the study. An example of this comparison is shown in FIG. 9. It was found that the logging detections overlapped with PRODES deforestation maps only 6% (±5%) of the time in any given year. Up to three years following harvest, a maximum of 19% (±11%) of logged areas were subsequently deforested (clear cut). Therefore, these results are not redundant with deforestation and thus represent forest damage that has been accounted for in previous State- and Basin-scale forest disturbance estimates.

The study covered the period 1999 to 2002, which is prior to the failure of the Scan Line Corrector (SLC) in the ETM+ instrument onboard Landsat 7. Following the SLC failure, roughly 40% of each acquired Landsat image is missing data. To seek out alternatives to Landsat 7, a satellite inter-comparison of detection capability based on the network bance estimates.

By means of the Monitoring the Brazilian Amazon Gross Deforestation (Terra-MODIS, AVHRR) data. The only sensor to meet or exceed the performance of Landsat ETM+ was EO-1 Hyperion; all others failed to detect at least 80% of the logging damage in the field sites. However, EO-1 Hyperion is a hyperspectral technology demonstration with extremely limited spatial and temporal coverage, making its application to large-area analysis intractable. It was concluded that the combination of spectral resolution (6 optical channels), spectral signal-to-noise performance, and spatial resolution was critically important in determining the amount of logging that could be reliably detected. Landsat 7 ETM+ provides the minimum performance needed for reliable analysis of selective logging in Amazon forests, and improved space-based technology is critically needed to remedy the current limitations.

Unobserved Areas

In some cases, there is a lack of sufficiently cloud-free imagery to determine logging extent using the pattern recognition portion of CLAS. For example, cloud cover precluded the use of two image pairs (Landsat path-row 222/062 and 226/062) in the state of Pará, one in 1999 and the other in 2000. To estimate logging in those areas, it was necessary to employ the single-scene analysis approach detailed by Asner et al. (2004a), which demonstrated that manual interpretation of single-date AutoMCU results have a temporal sensitivity of about one year following logging. Therefore, the extent of new logging in these two Pará images as a proxy for annual logging rates in these regions was estimated.

In other cases, persistent cloud cover resulted in no images for certain areas for certain years. In these instances, the amount of logging for the area was estimated from Landsat observations made in the closest year. Of the 480 Landsat images employed throughout the entire study, this was necessary 10.5% of the scenes.

Uncertainty Analyses

Any large-scale, complex remote sensing study must track and manage sources of uncertainty in the final results. This is important because there are many steps that can lead to errors. The uncertainty was carefully quantified in four key areas: (i) atmospheric correction (aerosol and water vapor), (ii) unobserved areas caused by persistent cloud cover, (iii) annualization, and (iv) auditor uncertainty.

Atmospheric Uncertainty

In the CLAS processing stream, Landsat ETM+ images are atmospherically corrected using the 6S atmospheric correction algorithm (Vermote, 1997), with monthly averages of aerosol and water vapor inputs from the MODIS satellite sensor. The sensitivity of the CLAS AutoMCU algorithm to atmospheric correction errors was comprehensively assessed by Asner et al. (2005), and was found to be minimally sensitive to uncertainties in aerosol and water vapor from MODIS. To further understand the effect that the atmospheric correction has on the sensitivity of entire CLAS process, five Landsat image pairs were atmospherically corrected using randomly-selected, monthly aerosol and water vapor values from MODIS. The difference in the amount of automatically detected logging between the different atmospherically-corrected images was only 0.7% (Table 2).

<table>
<thead>
<tr>
<th>Source</th>
<th>Percentage of Total Logged Area Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmosphere</td>
<td>±0.7%</td>
</tr>
</tbody>
</table>

Table 2: Primary sources of uncertainty in CLAS analyses of selective logging extent in forests of Amazonia.
Primary sources of uncertainty in CLAS analyses of selective logging extent in forests of Amazonia.

<table>
<thead>
<tr>
<th>Source</th>
<th>Percentage of Total Logged Area</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unobserved Area</td>
<td>±5%</td>
<td>Percentage of cloud- and shadow-covered area compared to total logged area</td>
</tr>
<tr>
<td>Annualization</td>
<td>±2-9%</td>
<td>Standard error of the difference between dry season length for matched pairs of consecutive years from 1979-1996.</td>
</tr>
<tr>
<td>Auditor</td>
<td>±12.8%</td>
<td>Standard error of difference between auditor estimates, on a per km² of logging basis</td>
</tr>
</tbody>
</table>

Total Estimated Error = ±11-14% Root mean square error

Unobserved Area Uncertainty
When cloud and cloud-shadow cover is greater than 50% in any 5,625 km² area (2,500x2,500 pixels), the area of observed logging is used to estimate the amount of logging in the unobserved, cloudy areas. The sensitivity to this type of error was assessed by simply quantifying the fractional cover of clouds and cloud shadows in comparison to observed logging extent. The calculated absolute uncertainty caused by this step was approximately ±5% over the five states (Table 2).

Annualization Uncertainty
Although the rate of logging is assumed constant throughout the dry season, there is a level of uncertainty inherent in this assumption. Marengo et al. reported rainy season length for five regions of the Amazon (e.g., North Amazonia, Central Amazonia, Mouth of Amazon, Southeast Amazonia, and Southwest Amazonia) for the period 1979-1996. To determine the uncertainty in the logging estimate related to assumption of dry season length, a series of matched pairs of dry season length for two consecutive years (e.g., 1979-1980, 1980-1981, 1995-1996) was compiled to calculate the standard error of the difference in dry season length for each region. This standard error (in days) was divided by the average length of the dry season for the respective region to express the uncertainty in percent of dry season. This percentage uncertainty was then applied to actual satellite image pairs or by averaging the uncertainty for states located between two regions. These uncertainties ranged from 2-9% as a result of interannual variation in dry season length (Table 2).

To further assess the sensitivity of the logging area estimates to the annualization and timing of the dry season, the estimates were also annualized without the constraint that logging activity only occurs only during the dry season. These results are reported in Table 3, with comparison to the preferred results that appear in the main text (Table 2). It is clear that the differences between these two assumptions can be large in the smaller states (e.g., Acre, Rondonia), where the estimate of logged areas is more sensitive to the acquisition dates of a smaller number of annual satellite image pairs. However, in the larger states, these uncertainties tend to balance out. In the majority of cases, the amounts of logging estimated without the dry season constraint still fall within the minimum and maximum limits of estimated logged area caused by other sources of uncertainty (discussed below).

**Table 3**

<table>
<thead>
<tr>
<th>Source</th>
<th>1999-2000 rates (km² yr⁻¹)</th>
<th>2000-2001 rates (km² yr⁻¹)</th>
<th>2001-2002 rates (km² yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry-</td>
<td>Calendar</td>
<td>Dry-</td>
<td>Calendar</td>
</tr>
<tr>
<td>season Protocol</td>
<td>Year</td>
<td>season Protocol</td>
<td>Year</td>
</tr>
<tr>
<td>Acre</td>
<td>64</td>
<td>91</td>
<td>53</td>
</tr>
<tr>
<td>Mato</td>
<td>13,843</td>
<td>11,702</td>
<td>7,912</td>
</tr>
<tr>
<td>Grosso</td>
<td>5,039</td>
<td>5,030</td>
<td>5,543</td>
</tr>
<tr>
<td>Rondonia</td>
<td>773</td>
<td>694</td>
<td>923</td>
</tr>
<tr>
<td>Roraima</td>
<td>32</td>
<td>32</td>
<td>55</td>
</tr>
<tr>
<td>TOTAL</td>
<td>20,651</td>
<td>17,609</td>
<td>14,286</td>
</tr>
</tbody>
</table>

Auditor Uncertainty
Each auditor reviewed a set of the same 25 image subsets (400 by 400 pixels) in which most images include some form of logging. A test was performed in which a novice and an experienced image analyst manually delineated areas containing logged forest. This comparison was used to calculate one standard error of the difference in logging assessments between auditors for each image subset. The standard error between auditors was 0.69 km² of logging, which when scaled by the average amount of logging identified by the two analysts (5.4 km²), resulted in an uncertainty of 12.8% (Table 2).

These different sources of uncertainty were compiled and used to estimate an overall uncertainty in the logging extent estimates of 11-14% for each Brazilian state in each year of analysis (Table 2). These uncertainties were then propagated to the Basin scale for annual estimates of selective logging for the years 2000, 2001, and 2002 (Table 4).

**Table 4**

<table>
<thead>
<tr>
<th>Source</th>
<th>1999-2000 rates (km² yr⁻¹)</th>
<th>2000-01 rates (km² yr⁻¹)</th>
<th>2001-02 rates (km² yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logged</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acre</td>
<td>54</td>
<td>78</td>
<td>45</td>
</tr>
<tr>
<td>Mato Grosso</td>
<td>11,801</td>
<td>16,521</td>
<td>6,744</td>
</tr>
<tr>
<td>Logged</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mato Grosso</td>
<td>11,801</td>
<td>16,521</td>
<td>6,744</td>
</tr>
</tbody>
</table>
Validation of CLAS logging detection method.

<table>
<thead>
<tr>
<th>Logging Block</th>
<th>Harvest Type</th>
<th>Image Date</th>
<th>Harvest Date</th>
<th>% Logging Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junauana</td>
<td>CL</td>
<td>2000</td>
<td>2001</td>
<td>16</td>
</tr>
<tr>
<td>Junauana</td>
<td>RIL</td>
<td>2000</td>
<td>2002</td>
<td>9</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2000</td>
<td>1997</td>
<td>0</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2000</td>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2000</td>
<td>2000</td>
<td>8</td>
</tr>
<tr>
<td>Junauana</td>
<td>RIL</td>
<td>2001</td>
<td>2002</td>
<td>1</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2001</td>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>CL</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>2001</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>1998</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>2001</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2000</td>
<td>2000</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>1999</td>
<td>4</td>
</tr>
<tr>
<td>Junauana</td>
<td>RIL</td>
<td>2001</td>
<td>2001</td>
<td>0</td>
</tr>
<tr>
<td>Junauana</td>
<td>CL</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2001</td>
<td>2000</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2002</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2002</td>
<td>2000</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2002</td>
<td>2000</td>
<td>16</td>
</tr>
<tr>
<td>Canaxi</td>
<td>CL</td>
<td>2000</td>
<td>1999</td>
<td>75</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2000</td>
<td>1999</td>
<td>59</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2000</td>
<td>1999</td>
<td>91</td>
</tr>
<tr>
<td>Tapajos</td>
<td>RIL</td>
<td>2000</td>
<td>2000</td>
<td>84</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2001</td>
<td>2000</td>
<td>44</td>
</tr>
<tr>
<td>Canaxi</td>
<td>RIL</td>
<td>2002</td>
<td>2002</td>
<td>36</td>
</tr>
</tbody>
</table>

Geographic information system (GIS) coverages of the logging areas listed in Table 5 were overlaid on the CLAS products, and statistical data on logging detection percentage area and logging type were calculated. Results were organized by success or failure in detecting the logged areas and their spatial extent. For analysis purposes, a false-negative detection was declared when CLAS missed areas logged in the 12 months prior to image date. A false-positive detection was declared for areas not logged in the 12 months prior to image date and when more than 25% of the false area was detected as logging. The 12-month limit, a result of the acquisition dates of the before and after AutoMCU images used in the change detection process, was selected based on the known sensitivity of the AutoMCU algorithm within CLAS (S3).

Of the 45 image combinations tested, only two false-positives and two false-negatives occurred (Table 5). For the two false-positives, the detected logging areas ranged from 34-38% of the true area and were logging blocks that had been harvested two years prior to satellite imaging. Therefore, it is considered that these two blocks as false-positives only in the sense that it was not intended for CLAS to find logging sites that are more than a year old. Further review of the false-positive from Canaxi pointed to issues of geo-registration of exact harvest boundaries between the GIS and the imagery. The other false-positive (Junauana) was very clearly re-harvested over roughly 40-50% of the originally delineated log-
ling block (which had been harvested three years prior to re-harvest), hence it was not considered further. For the two false-negatives, the detected logging areas ranged from 36-44% of the true area. Further review of the false-negative from Cauaxi also revealed co-registration problems between field data and remote sensing imagery. The other false-negative (Tapajós) occurred in a large block where the harvesting took place for several months, which caused the first harvested portions to be too old (14 months) for detection due to regrowth. 

A chi-square test of the results in Table 4 shows that there is a highly significant association between logging and CLAS algorithm detection ($\chi^2=17.0; p<0.001$ (Table 6)). This test is conservative because it is mainly testing the sensitivity to currently versus previously logged areas, and not to intact forest that have not been harvested.

**TABLE 6**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Logged No</th>
<th>Logged Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detected</td>
<td>37</td>
<td>2</td>
</tr>
<tr>
<td>Detected</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Expected</td>
<td>33.8</td>
<td>5.2</td>
</tr>
<tr>
<td>Detected</td>
<td>5.2</td>
<td>0.8</td>
</tr>
</tbody>
</table>

*Number of harvest/image combinations that met the following criteria: harvest months were known, when knowledge was essential for validation; detection was not obviously interfered with because of proximity of clouds/cumulus shadows.

Roundwood Estimates from Remote Sensing and Field Studies

The roundwood production was calculated from logging areas detected using the CLAS approach, combined with field-based relationships between remotely sensed canopy opening (PV cover), forest canopy gap fraction and roundwood harvest volumes. After the logging areas and canopy damage intensities were mapped as described in previous sections, an equation was applied to convert remotely sensed PV change of logged areas to forest canopy gap fraction. The equation was derived from intensive measurements of forest gap fraction in reduced-impact (low damage) and conventional (high damage) selective logging areas and co-located with Landsat ETM+ satellite imagery processed with the AutoMCU algorithm. This general conversion from the CLAS-derived PV fractional cover to forest canopy gap percentage was reported by Asner et al. (2005) as:

\[
\text{Canopy gap} = (\text{PV}_{\text{CLAS}} - 90.0) / (-0.4) \quad (2)
\]

\[
r^2=0.87, p<0.01
\]

where $\text{PV}_{\text{CLAS}}$ and canopy gap are in percentage units. $\text{PV}_{\text{CLAS}}$ is a planar metric, whereas canopy gap is the hemispherical canopy opening (SB). The gap-transformed data were then used to estimate the volume of roundwood ($m^3$) extracted on a per-area basis using an equation drawn from 35 logging sites in Brazil, Belize, Suriname, Guyana, and Indonesia (Pereira et al., 2002):

\[
\text{Wood volume} = 3.882 + 108.7 \times (\text{Canopy gap}/100) \quad (3)
\]

\[
r^2=0.83, p<0.0001
\]

Calculated roundwood extraction volumes were then compiled by logging detections (from CLAS), and mean harvest intensities were calculated by dividing the total calculated annual roundwood volume by the harvest area.

REFERENCES


Instituto-Socioambiental. (Sao Paulo, Brazil, 1999) Map of forest types, land-use change and protected areas in the Amazon.


What is claimed:

1. A method, comprising:
   - receiving satellite image data at a computer;
   - performing atmospheric correction on the satellite image data;
   - masking one or more of clouds, water, and deforestation in the satellite image data;
   - determining fractional material cover at least one of photosynthetic vegetation (PV) canopy, non-photosynthetic vegetation (NPV) canopy, and bare substrate;
   - generating sub-pixel fractional material cover images based on the fractional material cover of at least one of the PV canopy, the NPV canopy, and the bare substrate;
   - performing image differencing to create images of PV and/or NPV change; and
   - identifying areas of forest disturbance based on the images of PV and/or NPV change using pattern recognition.

2. The method of claim 1, wherein determining fractional material cover is performed at a sub-pixel scale from the masked satellite image data.

3. The method of claim 1, wherein the image differencing is performed on the sub-pixel fractional material.

4. The method of claim 1, wherein the pattern recognition includes pattern recognition of at least one of the following: logging decks, logging roads, skid trails, and tree felling gaps.

5. The method of claim 1, wherein the step of identifying further comprises analyzing a cluster of pixels using a moving kernel approach.

6. The method of claim 3, wherein the moving kernel approach includes positioning an analysis kernel at a starting base kernel with four subset kernels in each of the four corners of the base kernel.

7. The method of claim 6 further comprising the step of performing the analysis kernel to its neighboring kernel in at least one of the following directions—north, south, east, and west.

8. The method of claim 7 wherein the step of moving is performed if an area of the base kernel tests positive.

9. The method of claim 7 wherein the step of moving is performed until all neighboring kernels have been evaluated.

10. The method of claim 1 further comprising the step of performing an audit of probable logging event locations.

11. The method of claim 10 wherein the audit is performed to remove potential false logging event locations.

12. The method of claim 1 wherein the forest disturbance comprises spectrally decomposing each pixel of the masked satellite image data.

13. The method of claim 12 wherein the step of spectrally decomposing is performed using a probabilistic spectral sub-model.

14. The method of claim 13, wherein the step of spectrally decomposing is performed using a probabilistic spectral sub-model.

15. A system, comprising:
   - a computer including one or more processors and memory, the memory storing instructions for execution by the one or more processors, the instructions comprising:
     - performing atmospheric correction on the satellite image data;
     - masking one or more of clouds, water, and deforestation in the satellite image data;
     - determining fractional material cover of at least one of photosynthetic vegetation (PV) canopy, non-photosynthetic vegetation (NPV) canopy, and bare substrate;
     - generating sub-pixel fractional material cover images based on the fractional material cover of at least one of the PV canopy, the NPV canopy, and the bare substrate;
     - performing image differencing to create images of PV and/or NPV change; and
     - identifying areas of forest disturbance based on the images of PV and/or NPV change using pattern recognition.

16. The system of claim 15, wherein the instructions further comprise determining fractional material cover at a sub-pixel scale from the masked satellite image data.

17. The system of claim 16, wherein the instructions further comprise image differencing on the sub-pixel fractional material.

* * * * *