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.
EM 1899-2000 Logging

2000-2001 logging

2001-2002 Logging

Federal Conservation Units

Indigenous Reserves

Figure 1
Figure 2
Figure 5
REMOTE SENSING ANALYSIS OF FOREST DISTURBANCES

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was supported in part by funds obtained from NASA’s Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA-ECO), grant number NCC3-675 (I.C.-21). The U.S. Government may therefore have certain rights in the invention.

BACKGROUND OF THE INVENTION

Tropical forests have been threatened by increasing rates of deforestation or clear-cutting during the past three or more decades (E. F. Lambin, H. J. Geist, E. Lepers, Ann. Rev. Environ. Res. 28, 205 (2003)). Although deforestation, largely for conversion of land to food crops or pastures, is the major destructive force in tropical forests worldwide, other forest disturbances such as the selective harvest of timber have increased in frequency and extent (D. C. Nepstad et al., Nature 398, 505 (1999), L. M. Curran et al., Science 303, 1000 (2004)). In selective logging, a limited number of marketable tree species are cut, and logs are transported off-site to sawmills. Unlike deforestation that is readily observed from satellites, selective logging in the Brazilian Amazon causes a spatially diffuse thinning of large trees that is hard to monitor using satellite observations. Selective logging causes widespread collateral damage to remaining trees, sub-canopy vegetation and soils, with impacts on hydrological processes, erosion, fire, carbon storage, and plant and animal species.

There is surprisingly little known about the extent or impacts of selective logging throughout the tropical forests of the world, including the Amazon Basin. A survey of sawmills in the Brazilian Amazon suggested that 9,000-15,000 km² of forest had been logged in 1986-97 (D. C. Nepstad et al., Nature 398, 505 (1999)). The large uncertainty in this reported area resulted from necessary assumptions of the wood volume harvested per area of forest. Sawmill surveys can, at best, provide only a general idea of where and how much logging occurs because most operators buy timber at the mill gate rather than harvesting the wood themselves.

Objective, spatially-explicit reporting on selective logging requires either labor-intensive field surveys in frontier and often violently contested areas, or by remote detection and monitoring approaches. Previous studies of small areas show 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. 2 depicts a geographic coverage of study, showing the Brazilian Legal Amazon with Landsat 7 satellite footprints. FIG. 3 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. 4 depicts an example of deforestation and water body masking using Landsat thermal band 6 and the AutoMCU result for photosynthetic vegetation (PV).

FIG. 5 depicts a geographic coverage of study, showing the Brazilian Legal Amazon with Landsat 7 satellite footprints. FIG. 6 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. 7 depicts an example of deforestation and water body masking using Landsat thermal band 6 and the AutoMCU result for photosynthetic vegetation (PV).

FIG. 8 depicts a geographic coverage of study, showing the Brazilian Legal Amazon with Landsat 7 satellite footprints. FIG. 9 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. 10 depicts a block diagram of the CLAS system.

DETAILED DESCRIPTION OF THE INVENTION

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.

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. However, 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.
Conservation units such as indigenous reserves, parks and national forests generally afforded protection against logging. However, exceptions included areas in northern Mato Grosso, where up to 880, 291, and 50 km² of logging were measured each year in the Xingu, Aripuana, and Serra da Capivara, respectively (FIG. 1). In the southern portion of Pará state, major logging disturbances were observed in the Menkragnoti and Kayapó indigenous reserves, with up to 261 and 198 km² detected each year between 1999 and 2002. Federal forest reserves of Acre, Gorotire (Pará), and Jurumá (Mato Grosso) were harvested for timber at rates of up to 23, 90, and 380 km² each year, respectively.

Extensive field validation studies showed that the canopy damage detection within CLAS is precise and accurate, as set forth below in the Materials and Methods section. Field validation studies showed false-positive and false-negative detection rates of only 5%. Uncertainty caused by errors in atmospheric correction of satellite data, cloud cover, annualization, automated logging area delineation and manual auditing were 0.7-12.8% individually. After combining all known sources of error, the analysis suggests an overall absolute uncertainty of up to 14% in total logging area.

Selective logging contributes substantially to gross carbon fluxes from the Brazilian Amazon. Forest damage results from CLAS were combined with field-based forest canopy gap fraction and roundwood extraction data to calculate the total wood extraction rates. In 2000, 2001 and 2002, roundwood production averaged 49.8, 29.8 and 26.6 million cubic meters, respectively. The mean annual harvest intensities were 26.6, 21.7 and 21.4 m³ ha⁻¹, which were generally lower than those reported by sawmill owners in 1996. Nepstad et al. (1999) interviewed sawmill operators to estimate harvest intensities of 19, 28 and 40 m³ per hectare in 1996. The total volume harvested equates to 10-15 million metric tons of carbon removed. In addition to roundwood, residual stumps, branches, foliage and roots are left to decompose in the forest, subsequently returning to the atmosphere as carbon dioxide over about a decade. The calculated average harvest intensity of 23.2 m³ ha⁻¹ equates to ~8 Mg C ha⁻¹ contained in roundwood, with an associated 34-50 Mg C ha⁻¹ of fine and coarse debris. The conversion of roundwood to carbon assumes an average wood specific gravity of 0.7 Mg m⁻³ and a proportional carbon content of 0.5 as in Keller et al. (2004a). Fallen debris creation was estimated based on data from Keller et al. (2004b) based on mean debris amounts found in logged forests (~30 m³ ha⁻¹ harvested) subtracting the woody debris found in undisturbed forests. Upper and lower estimates were based on mean debris amounts plus root mean square (RMS) error accounting for the uncertainty of estimates for both background and logged sites. Total debris was estimated as 1.4 times fallen debris to account for standing dead and roots (Keller et al., 2001). Integrated to the regional scale, the processing of roundwood and decomposition of residues leads ultimately to a gross flux of carbon from the forest of up to 0.08 billion metric tons for each year of logging. The regional gross flux of carbon was estimated by multiplication of the range of carbon densities of debris created by the area logged. The range includes both variation in the annual area logged and uncertainty in the amount of debris created during log-
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 as being limited to these examples but rather should be construed as including 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 as being limited to these examples but rather should be construed as including in any way the remainder of the disclosure.

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 multiprocessor 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 here 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 (PV) 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 30x30 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 endmember "bundles", derived from extensive field databases and satellite imagery, to decompose each image pixel using the following linear equation:

\[
\rho(\lambda)_{pixel} = \sum C_{p}(\lambda)_{sub} \rho(\lambda)_{sub} + \epsilon
\]

where \( \rho(\lambda)_{pixel} \) 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 \( \rho(\lambda)_{pixel} \) and \( \rho(\lambda)_{sub} \) 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 (\( \rho(\lambda)_{sub} \)) 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), atmospheri-
periodically 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 shadowing, 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 linear 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 (Roberts 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 series of automated masks were 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 pixels 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 image 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 missed in the masks applied earlier in the CLAS process (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 7x7 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 6x6 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 toLogged Area Detection Procedure:

- Logging Node Map
- Thermal RMS mask (dynamically generated in earlier processing)
- PV mask (>60% fractional cover) (PV-mask)
- PV change difference image (PV 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 PV 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. The audit logging criteria are divided into high- and low-damage obvious and non-obvious categories. These categories encompassed all probable logging events in the study area and were the single-most ubiquitous surface feature found in harvested areas (Pereira et al., 2002; Asner et al., 2004). Abundance of logging decks

Obvious linear features (including primary-tertiary access roads and skid trails)

Severe canopy damage visible in PV change difference image

Area extent normally ?1 ha

Evidence of logging from previous years in close proximity (<15% PV change difference image)

High-Damage Non-Obvious Criteria

Few to no logging decks

Few to no linear features

Severe canopy damage visible in PV change difference image (>15%)

Presence of access roads or rivers, if not adjacent to an anthropogenic non-forest area

Area extent normally ?1 ha

Evidence of logging from previous years in close proximity (<15% PV change difference image)

Low-Damage Obvious Criteria

Few to no logging decks

Obvious linear features

Presence of access roads or rivers

Often tree-like in formation (graduating from higher to lower damage linear features)

Not a linear feature connecting non-forest areas, or otherwise used for general transportation

Evidence of logging from previous years in close proximity (<15% PV change difference image)

Low-Damage Non-Obvious Criteria

Few to no logging decks

Few to no linear features

Close proximity to access (i.e. roads, rivers or anthropogenic non-forest areas)

Speckles of recent canopy damage (felling gaps; >15%) in PV change difference image occurring at a density greater than in the surrounding forest areas

Area extent normally ?6.5 ha

Evidence of logging from previous years in close proximity (<15% PV change difference image)

Variation in the final logging products may result from differences in observer application of the manual audit procedure. The error associated with these user-specific differ-
The geographic overlap was evaluated between the Brazilian Space Research Institute’s deforestation maps 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 unaccounted 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 logging detection capability based on the network of low- and high-intensity logging sites in Amazonia was conducted (Asner et al., 2004). Comparisons were made among the detection capabilities of hyperspectral (EO-1 Hyperion), multi-view angle (Terra-MISR), high spatial resolution multi-spectral (EO-1 Advanced Land Imager, Landsat 5 Thematic Mapper, Landsat 7 ETM+, CBERS-2, SPOT, Terra-ASTER-VNIR), and low spatial resolution multi-spectral (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.

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 65 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 2

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>Atmosphere</td>
<td>±0.7%</td>
<td>Difference in automatically detected logged area between the atmospherically corrected image and an image with randomly selected atmospheric characteristics.</td>
</tr>
<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>
<tr>
<td>TOTAL ESTIMATED ERROR</td>
<td>±11-14%</td>
<td>Root mean square error</td>
</tr>
</tbody>
</table>

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 (i.e., 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 falls within the minimum and maximum limits of estimated logged area caused by other sources of uncertainty (discussed below).

TABLE 3

Logging estimates for Brazilian states in the Amazon using the dry-season annualization protocol from INPE (Marengo et al., 2001) and a calendar-year annualization.

<table>
<thead>
<tr>
<th>State</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></td>
<td>Dry-season Protocol Year</td>
<td>Dry-season Protocol Year</td>
<td>Dry-season Protocol Year</td>
</tr>
<tr>
<td></td>
<td>Dry-season Calendar Year</td>
<td>Calendar Year</td>
<td>Dry-season Calendar Year</td>
</tr>
<tr>
<td>Acre</td>
<td>64</td>
<td>91</td>
<td>53</td>
</tr>
<tr>
<td>Mato Grosso</td>
<td>13,843</td>
<td>11,762</td>
<td>7,912</td>
</tr>
<tr>
<td>Pará</td>
<td>5,939</td>
<td>5,030</td>
<td>5,343</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

Minimum-maximum logging estimates for Brazilian states in the Amazon based on uncertainties in CLAS logging methodology.

<table>
<thead>
<tr>
<th>State</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></td>
<td>Logged Minimum*</td>
<td>Logged Maximum†</td>
<td>Logged Minimum*</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>27,521</td>
<td>6,744</td>
</tr>
<tr>
<td>Pará</td>
<td>4,905</td>
<td>7,419</td>
<td>4,421</td>
</tr>
<tr>
<td>Rondônia</td>
<td>657</td>
<td>931</td>
<td>685</td>
</tr>
<tr>
<td>Roraima</td>
<td>27</td>
<td>38</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>17,444</td>
<td>24,987</td>
<td>12,041</td>
</tr>
</tbody>
</table>

*Composed of atmospheric, temporal interpolation, annualization, and auditor uncertainties (see text for definitions).
†Includes all uncertainties plus cloud interpolated area.
§Includes Northern Mato Grosso only.

### Validation

A comprehensive validation study of the logging extent results derived from the CLAS processing stream was carried out. Previous validation studies were highly detailed (tree-by-tree level) damage assessments, but were limited to fewer logging sites and only to the AutoMCU portion of the process (Asner et al., 2004a). Pereira et al., Asner et al. (2002b, 2004b), and Keller et al. (2004b) carried out extensive field studies from 1997-2002 in logging areas subjected to a wide range of harvest methods, intensities, and canopy damage levels. These studies included the development of high-resolution global positioning system (GPS) coverages of logging extent in conventional and reduced-impact logging sites in eastern Pará, central Pará, and northern Mato Grosso. These areas contained the most intensive and widespread logging in the entire study.

There were a total of 45 harvest/image combinations available for this validation study. The images areas were only considered where the harvest blocks were free of clouds and whose harvest month was known when that knowledge was essential. All logging events were contained in three Landsat images: Fazenda Cauaxi in eastern Pará (Landsat path/row 223/063); Tapajós National Forest in central Pará (path/row 227/062); and Juruena in northern Mato Grosso (path/row 229/067). The timber harvest dates of areas contained within these images ranged from 1997 to 2002 (Table 5). About half of the logging sites were harvested using conventional (high-damage) techniques, and the other half employed reduced-impact (low-damage) logging methods (Asner et al., 2004b). Logging areas ranged in size from 11 to 1,079 ha. This wide range of logging block sizes and canopy damage levels provided a substantial geographic data set against which to test CLAS.

### TABLE 5

Validation of CLAS logging detection method.

<table>
<thead>
<tr>
<th>Harvest Type</th>
<th>Image Date</th>
<th>Harvest Date</th>
<th>% Logging Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juruena</td>
<td>2000</td>
<td>2002</td>
<td>9</td>
</tr>
<tr>
<td>Tapajós</td>
<td>2000</td>
<td>1997</td>
<td>0</td>
</tr>
<tr>
<td>Tapajós</td>
<td>2000</td>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>Tapajós</td>
<td>2000</td>
<td>2000</td>
<td>8</td>
</tr>
<tr>
<td>Tapajós</td>
<td>2001</td>
<td>2002</td>
<td>1</td>
</tr>
<tr>
<td>Tapajós</td>
<td>2001</td>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1996</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1996</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1998</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>1999</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>2000</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>2000</td>
<td>0</td>
</tr>
<tr>
<td>Canaxi</td>
<td>2001</td>
<td>2000</td>
<td>36</td>
</tr>
<tr>
<td>Juruena</td>
<td>2002</td>
<td>2000</td>
<td>16</td>
</tr>
</tbody>
</table>

Harvests less than 12 months prior to satellite imaging.

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 differencing 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 Cauaxi pointed to issues of geo-registration of 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:

\[ \text{Canopy gap} = \frac{(P_{\text{CLAS}} - 90.0)}{(-0.4)} r^2=0.87, p<0.01 \]  

(2)

where \( P_{\text{CLAS}} \) and canopy gap are in percentage units. \( P_{\text{CLAS}} \) is a planar metric, whereas canopy gap is the hemispherical canopy opening (S8). 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):

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

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

(3)

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


an atmospheric correction module for correcting the image data of atmospheric effects;
an AutoMCU module for deconvoluting spectral signatures into sub-pixel fractional cover of live forest canopy, forest debris, and bare substrates;
an atmospheric adjustment module for cloud, water, and deforestation masking; and
a pattern recognition module having pattern recognition algorithms for forest disturbance mapping.

2. The system of claim 1, wherein the computer is a multi-processor system running on a Linux format.

3. An automated process to detect selective logging activity using spatial resolution data from sensors aboard orbiting bodies, the steps comprising:
obtaining orbital image data of regional and global events, correcting the data for atmospheric effects, spectrally decomposing spectral signatures into sub-pixel fractional cover of live forest canopy, forest debris, and bare substrates, masking the data for cloud, water, and deforestation, and applying pattern recognition algorithms for forest disturbance mapping.

4. The process of claim 3, wherein detection of selective logging activity is based on identified logging decks in the proximity of linear features and access with evidence of canopy damage and/or previous logging activity.

5. The process of claim 3, wherein probable logging activity is identified by analyzing a cluster of pixels using a moving kernel approach.

6. The process of claim 5, wherein the moving kernel approach includes positioning an analysis kernel at a starting base kernel with four subset kernels in each of four corners of the base kernel, and if an area of the base kernel tests positive, moving the analysis kernel to its neighboring kernel in a north, south, east, and west direction, and iterating the process until all neighboring kernels have been evaluated.

7. The process of claim 3, further including an auditing step for performing a manual visual audit of probably logging event locations to add or remove potential false positives and negatives.

8. The process of claim 3, wherein the process may be used to monitor any type of forest disturbance including from agriculture, cattle ranching, and natural hazards.

9. A computer program product comprising a non-transitory computer-readable medium having stored thereon computer executable instructions that, when executed by a computer, performs the process of claim 3.

* * * * *

References listed immediately above are hereby incorporated using various sensors aboard orbiting objects, wherein the logging activity using spatial resolution data obtained individually listed on the Provisional Application For Patent `software printouts and supporting documents, all of which are herein by reference in their entirety. This application includes section, and publication cited herein including but limited to the invention pertains and as may be applied to the invention following, in general, the principles of the invention intended to cover any variations, uses, or adaptations of the invention, and encompasses therefrom as modifications will be obvious to those skilled in the art.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as are obvious to those skilled in the art. The foregoing detailed description has been given for clearness of understanding only and no unnecessary limitations should be understood therefrom as modifications will be obvious to those skilled in the art.

The disclosures of each and every patent, patent application, and publication cited herein including but limited to the references listed immediately above are hereby incorporated herein by reference in their entirety. This application includes and incorporates in its entirety the attached CLAS software printouts and supporting documents, all of which are individually listed on the 'Provisional Application For Patent Cover Sheet' accompanying the filing of this application.

What is claimed:

1. An automated image processing system to detect selective logging activity using spatial resolution data obtained using various sensors aboard orbiting objects, wherein the system comprises:

a computer programmed to process orbital images of regional and global events, the computer including


The system comprises:

a computer programmed to process orbital images of regional and global events, the computer including