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
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 (LC-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 for large-scale selective logging assessments. A detailed comparison of Landsat satellite observations against field measurements of canopy damage following selective logging proved that traditional analytical methods missed about 50% of the canopy damage caused by timber harvest operations (G. P. Asner, M. Keller, R. Pereira, J. Zweede, Rem. Sens. Environ. 80, 483 (2002)).

BRIEF SUMMARY OF THE INVENTION

The present invention provides systems and methods for automatically analyzing Landsat satellite data of forests. The present invention can easily be used to monitor any type of forest disturbance, such as, but not limited to, logging, agriculture, cattle ranching, natural hazards (fire, wind events, storms), etc.
pattern recognition techniques. As discussed in greater detail
belongs. Although any methods and materials similar or
by one of ordinary skill in the art to which this invention
measured each year in the Xingu, Aripuana, and Serra
Acre
National Institute for Space Research (INPE) annual defor-
tive canopy openings, surface debris, and bare soil exposed by forest disturbances, and
approach provides automated image analysis using atmos-
the new Carnegie Landsat Analysis System (CLAS) to detect
matic Mapper Plus (ETM+) satellite data was advanced using
are included in the analysis.
Conservation units such as indigenous reserves, parks and
national forests generally afforded protection against log-
ners and materials similar or equivalent to those described herein can be used in the prac-
test of the present invention, the preferred methods
materials are described.
The computational analysis of Landsat Enhanced The-
cept described herein can be used in the practice or testing of the present invention, the preferred methods
The approach provides detailed measurements of forest
damage at a spatial resolution of 30x30 meters, and it
does so over millions of square kilometers of forest.
CLAS was applied to five states—Pará, Mato Grosso, Rond-
ian, Acre—that account for —90% of all deforestation in the Brazilian Amazon. The analysis was con-
ected on a time-series of Landsat ETM+ imagery from 1999
to 2002. Across the five timber producing Brazilian states, the
annual extent of selective logging ranged from 12,135 to
20.651 km² (FIG. 1). These logging results represent new
forest damage not accounted for in deforestation studies.
Each year, the overlap between the results and the Brazilian
National Institute for Space Research (INPE) annual defor-
ested Logan deforested Logan deforested
Selective logging was concentrated in the states of
Mato Grosso and Pará, where logging areas exceeded or
matched deforestation areas. In other smaller states,
selective logging increased forest damage area by 10-35%
over reported deforestation rates (Table 1).

| TABLE 1 |
| 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). |
| | 1999-2000 rates (km² yr⁻¹) | 2000-01 rates (km² yr⁻¹) | 2001-02 rates (km² yr⁻¹) |
| State | Logged | Deforested | Logged | Deforested | Logged |
| Acre | 64 | 547 | 53 | 419 | 111 | 727 |
| Mato Grosso | 13,843 | 6,176 | 7,912 | 7,504 | 7,267 | 6,880 |
| Pará | 5,939 | 6,671 | 5,343 | 5,237 | 3,791 | 8,697 |
| Rondônia | 773 | 2,465 | 923 | 2,673 | 946 | 3,605 |
| Roraima | 32 | 253 | 55 | 345 | 20 | 54 |
| Total | 20,651 | 16,112 | 14,286 | 16,178 | 12,135 | 19,963 |

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, round-
wood 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
interests of 19, 28 and 40 m² per hectare in 1996. The total
volume harvested equates to 10-15 million metric tons of
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 round-
wood, with an associated 34-50 Mg C ha⁻¹ of line and coarse
debris. The conversion of roundwood to carbon assumes an
average wood specific gravity of 0.7 Mg m⁻³ and a propor-
tional 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 for-
ests (~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 squared (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 lead
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-
ing. This value increases the estimated gross annual anthropo-
genic flux of carbon from Amazon forests by up to 25%
over carbon losses from deforestation alone. Post-harvest
forest regeneration reduces the net flux of carbon to the atmo-
sphere below these values but the pace of regeneration after
logging varies considerably.
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 ecol-
ology 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 obser-
vations and new computational techniques. The results, pre-
sented 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 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 clear-cutting 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 as being limited to these examples but rather should 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 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 (WV) values from the Moderate Resolution Imaging Spectrometer (MODIS) sensor onboard the NASA Terra spacecraft. Time-stamping of MODIS AOT and WV data alongside 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 end-member "bundles", derived from extensive field databases and satellite imagery, to decompose each image pixel using the following linear equation:

\[ \rho_{\lambda,\text{pixel}} = \sum_{k=1}^{3} (C_p, \rho_{k,\text{pixel}}) + C_{\text{sub-pixel}} \rho_{\text{sub-pixel}} + \epsilon \]

where \( \rho_{\lambda,\text{pixel}} \) 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 \( (\rho_{\lambda,\text{pixel}}) \) 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 \( (\rho_p(\lambda), \rho_{NPV}(\lambda), \text{and} \rho_{\text{sub-pixel}}(\lambda))) \) that encompass 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 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., 2005b).

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-
processing step, normalization of the forest values across
bances related to logging or other anthropogenic activities
by calculating the average change in fractional forest cover in
raw imagery before processing through the AutoMCU sub-
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 space-
borne 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 process-
ing (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 contain-
ing 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 thresh-
hold 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 corrections were 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 sub-
sets 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 distur-
bances 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 discriminat-
ing true disturbances from the surrounding forested areas.

Pattern Recognition

The specific criteria used in this procedure were deter-
mioned following a comprehensive analysis and review of the
forest responses to logging at various intensities in the Bra-
zilian states of Pará, Mato Grosso and Acre where field stud-
ies were conducted. The mean and standard deviation frac-
tional cover images from the AutoMCU step in CLAS
provide quantitative data on canopy damage and forest dis-
turbance intensity from which selectively logged areas can be
determined (FIG. 3). By identifying areas of canopy distur-
bance that are arranged in specific spatial patterns, it is pos-
sible to detect logged areas on an automated basis. The pri-
mary method by which logging is detected is image differ-
encing, 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 indi-
cate 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 log-
ging 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 log-
ging 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 cen-
tered 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 pat-
terns (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 pixel (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 Cl and NPV Cl 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)
PV change difference image (PV Cl)
NPV change difference image (NPV Cl)
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 NPV Cl>29%
Mean PV Cl>9%
Mean NPV Cl>29%
Mean PV Cl standard deviation>33%
Mean NPV Cl standard deviation>46%
More than 6 pixels (0.54 ha) with PV Cl values>80%
More than 6 pixels (0.54 ha) with NPV Cl values<85%
Masked area<0.18 ha

Subset Kernel Criteria:
≥2 subsets with PV Cl<32% standard deviation
≥2 subsets with mean PV Cl<3% and ≥60%
≥2 subsets having ≥1 pixel (0.09 ha) with a PV Cl value≥80%
≥2 subsets with NPV Cl<46% standard deviation
≥2 subsets with mean NPV Cl<5% and ≥65%
≥2 subsets having ≥1 pixel (0.09 ha) with a NPV Cl 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 northern 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 Forest Inventory Program (PRODES) 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.

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 these 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</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>
</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.

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 2-continued

<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

### 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>Dry-season Protocol</th>
<th>Calendar Year</th>
<th>Dry-season Protocol</th>
<th>Calendar Year</th>
<th>Dry-season Protocol</th>
<th>Calendar Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acre</td>
<td>64</td>
<td>91</td>
<td>53</td>
<td>48</td>
<td>111</td>
<td>117</td>
</tr>
<tr>
<td>Mato</td>
<td>13,843</td>
<td>11,702</td>
<td>7,912</td>
<td>7,783</td>
<td>7,267</td>
<td>7,182</td>
</tr>
<tr>
<td>Paris</td>
<td>5,039</td>
<td>5,030</td>
<td>5,343</td>
<td>5,159</td>
<td>3,791</td>
<td>3,751</td>
</tr>
<tr>
<td>Rondonia</td>
<td>773</td>
<td>694</td>
<td>923</td>
<td>902</td>
<td>946</td>
<td>638</td>
</tr>
<tr>
<td>Roraima</td>
<td>32</td>
<td>32</td>
<td>55</td>
<td>55</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**TOTAL** 20,651 17,609 14,286 13,947 12,135 11,708

### 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>2000-2001 rates (km² yr⁻¹)</th>
<th>2001-2002 rates (km² yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acre</td>
<td>54</td>
<td>78</td>
<td>45</td>
</tr>
<tr>
<td>Mato</td>
<td>11,801</td>
<td>16,521</td>
<td>6,744</td>
</tr>
</tbody>
</table>

**TOTAL** 6,195 8,453
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 results derived from the CLAS processing stream was carried out. These studies included the development of high-resolution geographic information system (GIS) 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); Tapajos 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>Logging Block</th>
<th>Type</th>
<th>Harvest Date</th>
<th>Image Date</th>
<th>% Logging Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cauaxi CL</td>
<td>2000</td>
<td>1996</td>
<td>1996</td>
<td>0</td>
</tr>
<tr>
<td>Cauaxi RIL</td>
<td>2000</td>
<td>1995</td>
<td>1995</td>
<td>0</td>
</tr>
<tr>
<td>Cauaxi CL</td>
<td>1998</td>
<td>1998</td>
<td>1998</td>
<td>0</td>
</tr>
<tr>
<td>Juruena CL</td>
<td>1999</td>
<td>1999</td>
<td>1999</td>
<td>5</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 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 exact harvest boundaries between the GIS and the imagery. The other false-positive (Juruena) was very clearly re-harvested over roughly 40-50% of the originally delineated log-
ging 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 there is a highly significant association between logging and CLAS algorithm detection ($\chi^2=17.0; 10.8; p<0.001$ (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>Logged</th>
<th>Observed</th>
<th>Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>37</td>
<td>2</td>
</tr>
<tr>
<td>Yes</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Expected</td>
<td>33.8</td>
<td>5.2</td>
</tr>
</tbody>
</table>

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

**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:

$$r^2=0.87, p<0.01$$

where $PV_{CLAS}$ and canopy gap are in percentage units. $PV_{CLAS}$ 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):

**Wood volume**

$$r^2=0.83, p<0.0001$$

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.
1. A method, comprising:
   (a) receiving satellite image data at a computer;
   (b) performing atmospheric correction on the satellite image data;
   (c) masking one or more of clouds, water, and deforestation in the satellite image data;
   (d) determining fractional material cover at least one of photosynthetic vegetation (PV) canopy, non-photosynthetic vegetation (NPV) canopy, and bare substrate;
   (e) 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;
   (f) performing image differencing to create images of PV and/or NPV change; and
   (g) 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 moving 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 includes at least one of the following: agriculture, cattle ranching, and a natural hazard.

13. The method of claim 1, wherein the step of determining further fractional material cover comprises spectrally decomposing each pixel of the masked satellite image data.

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) a computer including one or more processors and memory;
   (b) the memory storing instructions for execution by the one or more processors, the instructions comprising:
      (i) receiving satellite image data at a computer;
      (ii) performing atmospheric correction on the satellite image data;
      (iii) masking one or more of clouds, water, and deforestation in the satellite image data;
      (iv) determining fractional material cover of at least one of photosynthetic vegetation (PV) canopy, non-photosynthetic vegetation (NPV) canopy, and bare substrate;
      (v) 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;
      (vi) performing image differencing to create images of PV and/or NPV change; and
      (vii) 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.