Welcome to the CLASlite User Guide

The Carnegie Landsat Analysis System – Lite (CLASlite) is a software package designed for highly automated identification of deforestation and forest degradation from remotely sensed satellite imagery. This software package, its algorithms, and any derivatives are fully protected under patents 8189877, 20090214084 and 20120288159-A1; International Classification G06K9/62.

This guide provides information on CLASlite’s scientific background, technical processes, outputs, potential use, and limitations.

We trust that CLASlite will contribute to your organization’s forest monitoring efforts. For further information and general inquiries about the CLASlite program, contact us by e-mailing claslite@carnegiescience.edu or visit our website at http://claslite.ciw.edu.

CLASlite is made possible by the Gordon and Betty Moore Foundation and the endowment of the Carnegie Institution for Science.

Table of Contents

What's new in CLASlite v3.0.................................................................................................................................................. 3
Mission & Vision.................................................................................................................................................................. 4
The CLASlite Team............................................................................................................................................................ 5
About CLASlite.................................................................................................................................................................. 6
Step 1: Calibration to Reflectance ..................................................................................................................................... 7
Step 2: AutoMCU for Fractional Cover ............................................................................................................................. 10
Step 3: Forest Cover Classification .................................................................................................................................... 15
Step 4: Forest Change Detection ....................................................................................................................................... 17
Potential & Limits of CLASlite ........................................................................................................................................... 21
Useful References ............................................................................................................................................................... 24
Appendix I: Non-Disclosure Agreement............................................................................................................................ 27
Appendix II: Image Preparation for CLASlite..................................................................................................................... 28
Appendix III: Artifact removal using sliders.................................................................................................................... 30
WHAT'S NEW IN CLASlite v3.0

Version 3.0 of CLASlite has incorporated major breakthroughs in a variety of areas described throughout this manual. The new capabilities include:

- Extended spectral libraries for all tropical forests, from lowlands to mountain ecosystems.
- Improved deforestation and disturbance mapping algorithms.
- Faster spectral mixture analysis of satellite imagery supporting forest cover and change mapping.
- User-controlled image artifact removal to customize deforestation and disturbance mapping output.
- Batch processing of deforestation and disturbance images, up to 1000 entries at a time.
- New download manager that initiates CLASlite Setup. Setup process is more automated.
MISSION & VISION

MISSION

To expand the capacity of governments, non-governmental organizations, and academic institutions to map and monitor tropical forests with highly advanced and scientifically-based forest monitoring technology.

VISION

Following decades of scientific development, forest monitoring with satellites should be an everyday activity for non-experts, thus helping to improve environmental conservation, forest management, and resource policy development.

With CLASlite, governments, non-governmental organizations, and academic institutions of tropical forest nations will:

- Supplement their environmental toolkit with powerful, automated, user-friendly remote sensing technology.
- Map, monitor, and quantify the forest resource from a personal computer.
- Share information, best practices, and feedback within the CLASlite community, collectively and individually increasing forest-monitoring efficacy.

Community feedback based on using CLASlite in a variety of geographies and applications will guide further research and CLASlite technical development, continually advancing Carnegie technology to support forest-monitoring efforts of the CLASlite community.
THE CLASlite Team

The Carnegie Institution for Science (carnegiescience.edu) is a private organization that conducts basic research for the benefit of humanity. Carnegie’s Department of Global Ecology (globalecology.ciw.edu), located on the campus of Stanford University in California, is comprised of scientists conducting research on the interactions among the earth’s ecosystems, land, atmosphere, and oceans, with the goal of understanding the ways these interactions shape the behavior of the earth system, including its responses to future changes.

Gregory Asner is the Principal Investigator for the CLASlite Project. Greg is a faculty member in the Carnegie Institution’s Department of Global Ecology and Professor in Environmental Earth System Science at Stanford University.

Elif Tasar is the CLASlite Web Coordinator. Elif supports CLASlite training and dissemination via the CLASlite website.

Sinan Sousan is the CLASlite Programmer. Sinan programs the CLASlite processing environment and its graphical user interface, and he advances CLASlite technology by implementing user-driven improvements.

David Knapp is the Senior Remote Sensing Programmer. Dave develops new algorithms and data processing streams in support of CLAS, CLASlite and associated remote sensing systems.
ABOUT CLASlite

At a time when awareness of the role of forests in carbon storage, climate change mitigation, and biodiversity protection has dramatically increased, the Department of Global Ecology at the Carnegie Institution for Science seeks to rapidly advance the science of mapping forests to support internationally policy discussions, and to respond with applied solutions that address on-the-ground needs for forest monitoring.

The Carnegie Landsat Analysis System – Lite (CLASlite) is a software package designed for highly automated identification of deforestation and forest degradation from remotely sensed satellite imagery. Developed by Gregory Asner and his team at the Carnegie Institution, CLASlite incorporates state-of-the-art research in remote sensing into a simple, user-friendly yet powerful computer program intended for non-profit institutions and governments in need of technologies for forest monitoring and environmental planning.

CLASlite is the result of more than a decade of biophysical remote sensing research and fieldwork that provides an automated satellite mapping approach to determine one of the most important components of tropical forest structure: fractional cover of vegetation canopies, dead vegetation, and bare surfaces. These fractional covers are core determinants of ecosystem composition, physiology, structure, biomass, and biogeochemical processes. Fractional cover analysis sits at the heart of CLASlite, making it a powerful, stable and biophysically-grounded tool that allows for rapid forest monitoring with error tracking.

We at the CLASlite team have expanded our software capabilities and are capacity building for regional- and national-level forest monitoring. We are disseminating the technology through a tailored, demand-driven transference of CLASlite to government, academic and non-government (non-commercial) institutions based on available grant funding.

Currently, CLASlite supports input from seven different satellite sensors including Landsat 4 and 5 Thematic Mapper, Landsat 7 Thematic Mapper Plus, SPOT 4, SPOT 5, NASA ASTER, and NASA Advanced Land Imager (ALI) data. When Landsat 8 is ready in 2013, CLASlite v3.1 will be released to support it as well.

CLASlite Functions

CLASlite includes core functions to extract land cover information from raw satellite data, generating images and maps to support forest monitoring efforts. These processes include Calibration of raw imagery to apparent surface reflectance, Automated Monte Carlo Spectral Unmixing (AutoMCU) of reflectance data to fractional cover, classification of fractional cover data into a map of forest cover, and change detection with multi-temporal fractional cover data to map deforestation and forest disturbance. These functions are reflected in the steps of CLASlite’s User Interface, shown in Fig. 1.
STEP 1: CALIBRATION TO REFLECTANCE

OVERVIEW

All image sensors, from your eyes to your personal camera to an optical mapping satellite, are remote sensors. All optical imaging sensors are designed to measure the variation in color from pixel to pixel. Raw imagery can be calibrated and atmospherically corrected to reveal valuable surface reflectance information (i.e. color), a critical element in vegetation mapping.

SCIENTIFIC BACKGROUND AND TECHNICAL PROCESS

Radiometric Calibration

When a satellite-based sensor records data, it detects energy reflected from the land surface and the atmosphere between land and the sensor head. These data are collected by the imaging system for on-board digital storage and/or transmission to ground receiving stations.

To use an image quantitatively, however, the data registered in each pixel must be calibrated from units of digital numbers or counts, to units of reflected energy. This process is called radiometric calibration. For radiometric calibration, CLASlite uses conversion factors (gains and offsets) made available by the providers of satellite sensors (i.e. NASA, SPOT, etc.). The result of radiometric calibration is an image of in units of radiance (i.e. watts per square meter per unit of solid angle), also known as the energy measured by the satellite-based sensor.

Atmospheric Correction

Radiometric data contain information about both the Earth’s surface and its atmosphere. Thus, to work with vegetation (surface data), it is necessary to minimize the contribution of the atmosphere to the values of each pixel in the satellite image. This is accomplished through a process called atmospheric correction, which minimizes the effect of water vapor (humidity), aerosols (from dust, volcanoes, etc.), and other factors.

To apply atmospheric correction, CLASlite uses the 6S radiative transfer model (Vermote et al. 1997), which simulates the Earth’s atmosphere in each satellite image. Using data from NASA’s Moderate Resolution Imaging Spectroradiometer (MODIS) sensor, 6S models the effect of the atmosphere on sunlight as it passes through the atmosphere, interacts with the land surface, and returns through the atmosphere to the satellite sensor. The raw image is then “corrected” by removing the estimated model of the atmosphere, resulting in an image of surface reflectance (i.e. units of percentage, represented by integers 0 to 10000, where 10000 corresponds to 100%).

For every image, CLASlite has fully automated the integration of monthly averages of MODIS data, corresponding to the image’s acquisition date, into the 6S code.

Masking
Mapping with optical satellite sensors (Landsat, SPOT, etc.) requires radiance data to determine the reflectance of each pixel, which is the information required to extract information about vegetative cover. No satellite sensors can collect this radiance information on the land surface through clouds, in darkened shadow areas under clouds, or in shadows caused by steep terrain. Thus, clouds and cloud shadows, as well as terrain shadows must be masked, or excluded from the image analysis. These, as well as bodies of water, are automatically masked out of each image during the calibration step within CLASLite.

**OUTPUT**

The result of CLASLite image calibration is a reflectance image, providing spectral bands calibrated from raw data to apparent surface reflectance. The number of bands varies by sensor (e.g. Landsat: 6 reflectance bands; SPOT: 4 bands, etc.)

**Reflectance Image Band Analysis**

After an image has been converted to apparent surface reflectance, the image can be viewed with the ENVI Freelook or similar software to review the reflectance profile for each pixel. In Fig. 2, bands 5, 4, 3 (RGB) are displayed in this Landsat-7 example image.

If you place your cursor over a forested pixel in ENVI Freelook, you can see a spectral profile of that pixel by selecting Options → Z Profile.

In the spectral profile, the X axis represents the band number and the Y axis represents reflectance (% times 100). In this example image, there are 6 bands that represent six of the seven spectral bands provided by Landsat 7. (The seventh band, Band 6, is a thermal image.) You can see that the vegetation has a near-infrared (NIR, Band 4) reflectance of more than 30%, while the red reflectance (band 3) is only 2%. This difference between the NIR and red reflectance is characteristic of vibrant, green vegetation. The line between these two bands is often referred to as the “red edge”.

You can also see that in the visible bands (1, 2, and 3), this pixel is brighter in the green (Band 2) than in the blue and red (Bands 1 and 3, respectively). Vegetation absorbs more light in the blue and red parts of the spectrum, leaving green as the color we see with the naked eye.

![Figure 2: Pixel reflectance profile from a Landsat 7](image-url)
In a pixel that has little to no vegetation, you can see that the NIR reflectance is lower and the red reflectance is higher than in vegetated pixels. The typical form that we saw in the vegetation with peaks at bands 2 and 4 does not exist in the spectrum of this pixel, indicating that there is less vegetation. In addition, reflectance in bands 5 and 7 (the shortwave-infrared or SWIR) is much higher in pixels with less or no live, green vegetation. This indicates the presence of dead, brown vegetation, soils, and rock.
**STEP 2: AUTO MCU FOR FRACTIONAL COVER**

**OVERVIEW**

Different types of Earth surface covers have different reflectance properties. In other words, each component of the Earth surface has a spectral signature. From these spectral signatures, it is possible to derive information on each pixel in a reflectance image. Since the 1960s, it has been possible to generate maps of land cover from remotely sensed imagery using classification techniques, which assign a whole pixel to a class (i.e., forest, rock) based on the spectral signature in the pixel. This type of thematic classification is useful for land-cover mapping, but it often has reduced sensitivity to small variations and changes in forest cover that occur at the sub-pixel, or within-pixel, scale. Since we want to map deforestation and forest degradation occurring at the sub-pixel scale, we must use a different approach.

CLASlite is the result of more than a decade of biophysical remote sensing research and fieldwork that provides an automated satellite mapping approach to determine arguably the most important characteristic of any forest: the fractional cover of live vegetation canopy, dead vegetation, and bare surface within the forest ecosystem. These fractional covers are core determinants of forest composition, structure, biomass, physiology and biogeochemical processes. Fractional cover analysis lies at the heart of CLASlite, providing a powerful, stable and biophysically-grounded algorithm that allows rapid forest monitoring with error tracking.

**AUTO MCU – SCIENTIFIC BACKGROUND AND TECHNICAL PROCESS**

The AutoMCU, or Automated Monte Carlo Unmixing (Asner 1998, Asner and Heidebrecht 2002, Asner et al. 2004), provides quantitative analysis of the fractional or percentage cover (0-100%) of live and dead vegetation, and bare substrate within each satellite pixel (e.g., within each 30 x 30 m pixel in a Landsat image). Live vegetation is technically referred to as Photosynthetic Vegetation (PV) because live vegetation maintains unique spectral properties associated with leaf photosynthetic pigments, canopy water content, and the amount of foliage in the canopy. The dead or senescent vegetation fraction is termed Non-photosynthetic Vegetation (NPV), which is expressed in the spectrum as bright surface material with spectral features associated with dried carbon compounds in dead leaves and exposed wood. Finally, bare substrate is often dominated by exposed mineral soil, but can also be rocks and human-made infrastructure (e.g. brick).

The AutoMCU was initially developed for forest, savanna, woodland and shrubland ecosystems (Asner 1998, Asner and Lobell 2000, Asner and Heidebrecht 2002), and was later redesigned for tropical forests (Asner et al. 2004, 2005). The method requires “libraries” of spectral endmembers for each of three relevant surface cover types: bare substrate, photosynthetic vegetation, and non-photosynthetic vegetation. Endmembers are reference spectra that are chosen as pure representatives of a given surface material, and they are intended to encompass the spectral variability within that surface material. These libraries, derived from extensive field databases and satellite imagery, are used to decompose each image pixel using the following linear equation:

\[ \rho(\lambda)_{\text{pixel}} = \Sigma [C_e \cdot \rho(\lambda)_e] + \varepsilon = [C_{\text{pv}} \cdot \rho(\lambda)_{\text{pv}} + C_{\text{npv}} \cdot \rho(\lambda)_{\text{npv}} + C_{\text{substrate}} \cdot \rho(\lambda)_{\text{substrate}}] + \varepsilon \]  

(1)
where $\rho(\lambda)_e$ is the reflectance signature library (e) at wavelength $\lambda$, and $\varepsilon$ is an error term. Solving for each sub-pixel cover fraction ($C_e$) requires that the satellite observations ($\rho(\lambda)_{\text{pixel}}$) contain sufficient spectral information to solve a set of linear equations, each of the form in equation (1) but at different wavelengths ($\lambda$).

The tropical forest spectral libraries provide the spectral reflectance signatures required by the AutoMCU sub-model: $\rho_{\text{pv}}(\lambda)$, $\rho_{\text{npv}}(\lambda)$, and $\rho_{\text{substrate}}(\lambda)$. The AutoMCU is a probabilistic approach based on canopy physics (Asner 1998) that reduces each image pixel into the three constituent cover fractions of PV, NPV and bare substrate.

**The AutoMCU Spectral Libraries**

For the tropical forest spectral library used in CLASlite, both the bare substrate and NPV spectra were collected using ground-based field spectroradiometers (FR and FS-3 Analytical Spectral Devices, Inc., Boulder, Colorado USA). The bare substrate library incorporates a diverse range of mineral soil types, surface organic matter levels and moisture conditions. The NPV spectra library includes surface litter, senescent grass, deforestation residues (slash), and other dry carbon constituents collected from a wide range of species and decomposition states.

In contrast to bare substrate and NPV, the PV spectra of forest canopies require overhead viewing conditions, which is difficult with trees reaching heights of more than 50 meters. Spectral measurements of individual leaves, stacks of foliage, or partial canopies (e.g., branches) introduce major errors in spectral mixture models requiring canopy-level information (Asner 2008). To develop a canopy-level spectral library for CLASlite, PV spectral data were collected using the Earth Observing-1 (EO-1) Hyperion sensor (Ungar et al. 2003), which is the only spaceborne imaging spectrometer launched by NASA for environmental applications. Hyperion data were collected over many tropical forest control sites in Brazil, Peru and elsewhere from 1999 to 2012, providing many millions of spectral observations made at 30-m resolution (Asner et al. 2005, Asner 2008, Asner unpublished data). These hyperspectral data were atmospherically corrected to reflectance and convolved to the spectral channels used by the Landsat, ALI, ASTER, and SPOT sensors in CLASlite. As a result, these datasets incorporate the highly variable effects of intra- and inter-crown shadowing, common in tropical forests (Asner and Warner 2003). In total, the spectra represent more than 250,000 field and spaceborne spectrometer observations.

**Automated Monte Carlo Unmixing (AutoMCU)**

The AutoMCU iteratively selects a PV, NPV and bare substrate spectrum from each library, and unmixes the pixel reflectance into constituent cover fractions using equation (1). CLASlite adopts a Monte Carlo method, whereby the possible combinations of the endmember spectra are pre-computed, and are applied during the AutoMCU run. The process of random selection is repeated up to 50 times or until the solution converges on a mean value for each surface cover fraction. In the original CLAS [20], the iteration was done dynamically until a stable standard deviation between successive fractional cover estimates was reached. Following a series of studies on different tropical forests, we found that 50 iterations per pixel is usually sufficient to achieve a stable solution based on this Monte Carlo approach, and thus this value is fixed in CLASlite (Fig. 3).

An advantage of the Monte Carlo approach is that the per-pixel iterations produce a standard deviation of the estimate for PV, NPV and bare substrate fractions (Fig. 3). These are output from CLASlite as
standard deviation images. In addition, a final analysis of the fit of the modeled spectrum (right side of eq. 1) to the input spectrum (left side of eq. 1) is computed for each pixel, leading to a root mean squared error (RMSE) image.

Figure 3: Processing Stream for the Automated Monte Carlo Unmixing (AutoMCU) sub-model within CLASlite
OUTPUT

The output from AutoMCU in CLASlite is a 7-band image containing information about fractional cover of PV, NPV, and bare substrate, uncertainty estimates for each cover fraction, and total error for each pixel in the image.

Fractional Cover Image Bands

<table>
<thead>
<tr>
<th>Band</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1</td>
<td>Fractional cover of bare substrate (S), expressed as a percentage (0-100%)</td>
</tr>
<tr>
<td>Band 2</td>
<td>Fractional cover of photosynthetic vegetation (PV), expressed as a percentage (0-100%)</td>
</tr>
<tr>
<td>Band 3</td>
<td>Fractional cover of non-photosynthetic vegetation (NPV), expressed as a percentage (0-100%)</td>
</tr>
<tr>
<td>Band 4</td>
<td>Uncertainty of the S fraction, expressed as the standard deviation of AutoMCU iterations</td>
</tr>
<tr>
<td>Band 5</td>
<td>Uncertainty of the PV fraction, expressed as the standard deviation of AutoMCU iterations</td>
</tr>
<tr>
<td>Band 6</td>
<td>Uncertainty of the NPV fraction, expressed as the standard deviation of AutoMCU iterations</td>
</tr>
<tr>
<td>Band 7</td>
<td>Total error, expressed as the RMSE of the modeled versus observed reflectance signature</td>
</tr>
</tbody>
</table>

Figure 4: The fractional cover output from CLASlite

The fractional cover image can be analyzed visually, by displaying a color composite of bands 1-3 in an image viewer like ENVI Freelook. In Fig. 4, Band 1 (Fractional cover of S) is displayed in red, Band 2 (Fractional cover of PV) is displayed in green, and Band 3 (Fractional cover of NPV) is displayed in blue. The intensities of each color represent presence of each cover type in each pixel. For example, greener pixels have higher percentage of PV, yellow pixels indicate the presence of both S and PV, while bluer pixels represent higher fractional coverage of NPV.

The image can also be analyzed quantitatively through evaluation of the image band values. Bands 1-3 represent the fractional cover of PV, NPV, and S, expressed in percentages (0-100%). Because these fractions are the output of a probabilistic model, the percentages may not sum exactly to 100%. Bands 4-6 represent uncertainty of each cover fraction, expressed as the standard deviation (SD) of the 50 AutoMCU iterations for each cover type. Higher values indicate increased uncertainty. Band 7
represents the total error, or root mean squared error (RMSE) of the observed versus the modeled reflectance spectra. Combined, the standard deviation and RMSE images provide a way to assess the performance of the AutoMCU on a pixel-by-pixel basis, allowing you to identify areas of concern. Such areas can occur when a vegetation type is not well represented in the spectral libraries, in areas where inorganic materials are present (e.g., infrastructure), or atmospheric disturbances remain unmasked from other CLASlite steps (e.g., edges of clouds, severe haze).
**STEP 3: FOREST COVER CLASSIFICATION**

**OVERVIEW**

CLASlite can be used to map forest canopy cover from a single satellite image. The patterns and spatial orientation of forest and non-forest cover within these forest cover images often indicate areas of past clear-cutting and disturbance (e.g., an sharply delineated patch of non-forest within an otherwise forested swath), as well as natural non-forest vegetation such as grasslands and shrublands (e.g., an expansive non-forest area with natural borders).

**SCIENTIFIC BACKGROUND AND TECHNICAL PROCESS**

In version 3.0 of CLASlite, this simple decision tree is used to convert the single-image AutoMCU results to an estimate of forest cover.

Forest: \[ PV \geq 80 \text{ AND } S < 20 \]
Non-forest: \[ PV < 80 \text{ OR } S \geq 20 \]

where PV is photosynthetic vegetation cover fraction in the pixel, S is the bare substrate fraction in the pixel, and both thresholds can be customized, allowing you to tune your work to the current forest conditions. This S term is included to eliminate non-forest regrowth (successional vegetation, grasses, and some agriculture which may contain high PV fractions) from the forested class. These regrowth cover types usually have higher S levels than are found in neighboring intact forest.

This simple decision tree for forest cover, based on a default S setting of 20, is sufficiently general to allow the algorithm to accommodate a broad range of tropical forests. However, we recommend that you, the user, independently validate these maps. Based on validation, we encourage you to develop more accurate forest cover mapping in your area of interest by improving the forest cover criteria. This can be done by adjusting the Sval threshold value in CLASlite or by applying a user-modified decision tree to the fractional cover image outside of CLASlite. The power of CLASlite thus rests in quickly providing calibrated and corrected reflectance and fractional cover results, at which point you can make an informed modification to the forest cover mapping approach.

**OUTPUT**

The output of Step 3 in CLASlite is classified map of forest cover. The map contains three classes, defined below:

0 – Masked pixels
1 – Forest
2 – Non-forest

Figure 5: Fractional cover (left) and forest cover (right)
Forest cover maps can be ingested into third party geographic information systems (GIS) for the calculation of spatial statistics or conversion into visual and printable maps. While non-forest areas represented in the forest cover map may be the result of deforestation, it is important to note that neither deforestation nor disturbance can be mapped using a single satellite image. Deforestation and disturbance are forest change, and thus require multiple images for detection.
**STEP 4: FOREST CHANGE DETECTION**

**OVERVIEW**

CLASlite includes the fully automated capability to detect *forest change* between a time series of images taken of the same geographic area over time. Multi-image analysis is the most accurate approach for detection of forest loss (deforestation), gain (secondary regrowth), or degradation (areas of persistent forest disturbance).

**SCIENTIFIC BACKGROUND AND TECHNICAL PROCESS**

To map forest change, CLASlite applies the following decision trees to each pair of images, where the subscripts 1 and 2 indicate images from one year to the next.

Where:

- $PV_1 = 1^{st}$ Image photosynthetic vegetation fraction
- $NPV_1 = 1^{st}$ Image non-photosynthetic vegetation fraction
- $S_1 = 1^{st}$ image bare substrate fraction
- $RMSE_1 = 1^{st}$ image RMSE
- $PV_2 = 2^{nd}$ image photosynthetic vegetation fraction
- $NPV_2 = 2^{nd}$ Image non-photosynthetic vegetation fraction
- $S_2 = 2^{nd}$ image bare substrate fraction
- $RMSE_2 = 2^{nd}$ image RMSE
- $Refl_{1b1} = 1$st image reflectance band 1
- $Refl_{2b1} = 2$nd image reflectance band 1
- $Refl_{1b4} = 1$st image reflectance band 4
- $Refl_{2b4} = 2$nd image reflectance band 4

**a) Deforestation and disturbance pixels are calculated.**

**Deforestation:**

- $(((PV_1 - PV_2) \geq 25) \text{ [PV decrease captures most deforestation]}
- OR $((S_1 \leq 5) \text{ AND } ((S_2 - S_1) \geq 15)) \text{ [S increase captures deforestation followed by early regrowth]}
- OR $((PV_2 < 80) \text{ AND } ((NPV_2 - NPV_1) \geq 20)) \text{ [NPV increase]}

**Forest Disturbance:**

- $(((NPV_2-NPV_1) \geq 10) \text{ AND } ((PV_1-PV_2) > 10)) \text{ OR } ((S_1 \leq 5) \text{ AND } ((S_2-S_1) > 10) \text{ AND } (S_2 \leq 15))$

**b) The following pixels are excluded from the forest change analysis** [to eliminate the detection of false positives]:

For Landsat sensors:
For Deforestation:

\[
((PV_1 \leq 0) \text{ AND } (NPV_1 \leq 0) \text{ AND } (S_1 \leq 0)) \text{ [masked pixels (image 1)]} \\
\text{OR } ((PV_2 \leq 0) \text{ AND } (NPV_2 \leq 0) \text{ AND } (S_2 \leq 0)) \text{ [masked pixels (image 2)]} \\
\text{OR } ((PV_1 < 80) \text{ OR } (S_1 \geq 15)) \text{ [non-forest pixels (image 1)]} \\
\text{OR } ((PV_1 \geq 80) \text{ AND } (NPV_1 \geq 35) \text{ AND } (RMSE_1 \geq 6)) \text{ [unmasked cloud shadows and water (image 1)]} \\
\text{OR } ((PV_2 \geq 80) \text{ AND } (NPV_2 \geq 35) \text{ AND } (RMSE_2 \geq 6)) \text{ [unmasked cloud shadows and water (image 2)]} \\
\text{OR } ((S_2 \geq 50) \text{ AND } (S_1 < 100) \text{ AND } (PV_2 > 0)) \text{ [unmasked cloud rings (image 2)]} \\
\text{OR } (((NPV_2 - NPV_1) < 10) \text{ AND } (abs(Refl_{1b_1} - Refl_{2b_1}) > 300)) \text{ [unmasked cloud rings, cloud shadows, and topography mountain shadows]}
\]

For Disturbance:

\[
((PV_1 \leq 0) \text{ AND } (NPV_1 \leq 0) \text{ AND } (S_1 \leq 0)) \text{ [masked pixels (image 1)]} \\
\text{OR } ((PV_2 \leq 0) \text{ AND } (NPV_2 \leq 0) \text{ AND } (S_2 \leq 0)) \text{ [masked pixels (image 2)]} \\
\text{OR } ((PV_1 < 80) \text{ OR } (S_1 \geq 15)) \text{ [non-forest pixels (image 1)]} \\
\text{OR } ((PV_1 \geq 80) \text{ AND } (NPV_1 \geq 35) \text{ AND } (RMSE_1 \geq 6)) \text{ [unmasked cloud shadows and water (image 1)]} \\
\text{OR } ((PV_2 \geq 80) \text{ AND } (NPV_2 \geq 35) \text{ AND } (RMSE_2 \geq 6)) \text{ [unmasked cloud shadows and water (image 2)]} \\
\text{OR } ((S_2 \geq 50) \text{ AND } (S_1 < 100) \text{ AND } (PV_2 > 0)) \text{ [unmasked cloud rings (image 2)]} \\
\text{OR } (((NPV_2 - NPV_1) < 10) \text{ AND } (abs(Refl_{1b_1} - Refl_{2b_1}) > 300)) \text{ [unmasked cloud rings, cloud shadows, and topography mountain shadows]}
\]

For all other, non-Landsat, sensors (SPOT, ALI, and ASTER):

\[
\text{For Deforestation and Disturbance:} \\
\text{((PV_1 \leq 0) \text{ AND } (NPV_1 \leq 0) \text{ AND } (S_1 \leq 0)) \text{ [masked pixels (image 1)]} \\
\text{OR } ((PV_2 \leq 0) \text{ AND } (NPV_2 \leq 0) \text{ AND } (S_2 \leq 0)) \text{ [masked pixels (image 2)]} \\
\text{OR } ((PV_1 < 80) \text{ OR } (S_1 \geq 15)) \text{ [non-forest pixels (image 1)]} \\
\text{OR } ((PV_1 \geq 80) \text{ AND } (NPV_1 \geq 35) \text{ AND } (RMSE_1 \geq 6)) \text{ [unmasked cloud shadows and water (image 1)]} \\
\text{OR } ((PV_2 \geq 80) \text{ AND } (NPV_2 \geq 35) \text{ AND } (RMSE_2 \geq 6)) \text{ [unmasked cloud shadows and water (image 2)]} \\
\text{OR } ((S_2 \geq 50) \text{ AND } (S_1 < 100) \text{ AND } (PV_2 > 0)) \text{ [unmasked cloud rings (image 2)]}
\]

Pixels that meet the above criteria are excluded from both the deforestation and disturbance images. User customization of artifact removal thresholds is currently only supported for Landsat imagery.

c) Isolated pixels for deforestation and disturbance undergo filtration.

A spatial filter (each deforestation pixel must be surrounded by a minimum of 5 deforestation pixels within a 3x3 pixel kernel) is applied to the raw deforestation result to remove isolated pixels. Pixels that meet the deforestation criteria but do not pass this filter are added to the group of pixels classified as disturbance. An additional spatial filter (each disturbance pixel must be surrounded by a minimum 5 disturbance pixels within a 7x7 pixel kernel) is applied to the raw disturbance result to remove isolated pixels while conserving detected patterns often associated with forest disturbance. These filters are conservative, in that we employ them in CLASLite to reduce your sensitivity to natural tree-fall events and spurious artifacts. As a result, you may under-estimate deforestation and disturbance in some instances. The disturbance filter is less conservative than the deforestation filter because disturbance is more likely to occur in isolated patches.

Deforestation sliders for Landsat imagery
The forest change outputs may include unwanted artifacts (false positives) caused by the influence of clouds, unmasked cloud edges, cloud shadows, topography, and water boundaries. For Landsat imagery only, you can define desired thresholds for artifact removal in both the deforestation and disturbance images. These thresholds, which you can adjust by toggling the two artifact removal sliders in the step 4 window, determine the numerical values in the final criterion of the decision trees shown above. A slider setting of 0 percent means “no removal of artifacts” and a slider setting of 100 percent means “removal of all possible artifacts.” The default slider values, which correspond to the values shown above, are 50 percent for deforestation and 25 percent for disturbance. These defaults are intended to remove noise as best as possible without removing real forest change, such as pixels of grass cover on previously deforested land.

As you adjust the disturbance threshold, the recommended deforestation threshold will correspondingly increase or decrease. This relationship is due to a link between the process in step b) and c): because pixels determined by the deforestation filter to be artifacts are passed on to the disturbance filter to be considered as possible disturbance, a high deforestation slider value is generally needed to achieve a high disturbance slider value. This automatic adjustment is a recommendation only; you can and should ignore this suggestion if you know the study area well enough to estimate what percentage of artifact removal is appropriate.

For illustrated examples of the effectiveness and limitations of the sliders, see Appendix III.

**Aggregation feature**

By default, the deforestation and disturbance pixels undergo one final step in order to generate a deforestation map in which contiguous forest loss areas are clearly depicted. In this step, disturbance pixels in close proximity to a contiguous patch of deforestation pixels are moved from the disturbance output to the deforestation output. To achieve this, a buffer representing a radius of 120 meters is applied to the boundary of contiguous deforestation patches and all disturbance pixels falling within that radius is moved to the deforestation map. The outcome is that, for every deforestation patch, disturbance pixels associated with the same forest loss event will appear in the final deforestation output. The 120 meter buffer may be conservative in some places—that is, it may not capture all disturbance associated with a contiguous deforestation patch—so you should always check your results.

You can choose to map deforestation and disturbance without this Aggregation Feature by unselecting the “Aggregation Feature” box prior to running the change detection. You should recognize that turning off the aggregation feature discards some of the real-life spatial relationships between deforestation and nearby associated disturbance. For example, you might encounter a donut-shaped patch of deforestation with a couple disturbance pixels, in which only one tree was logged, embedded inside.

One situation in which we recommend that you turn the Aggregation Feature off is in areas prone to large-scale, natural disturbance. River corridors, for example, often undergo high levels of natural disturbance due to increases and decreases in river height as well as meandering of the riverbed. Other types of large-scale natural disturbance may be found in certain vegetation types, such as bamboo forests of the southwestern Amazon and Southeast Asia. In such cases, the Aggregation Feature will attempt to reclassify disturbances neighboring an actual deforestation event, thereby over-estimating the footprint of deforestation.

**Multi-image analysis**
If Step 4 is run with more than two images (up to 10), the results from individual time steps are compiled into a single map for each deforestation and forest disturbance for the entire time series according to one of the following methods, specified by you:

**First change:** The first detected events of deforestation and disturbance within the time series are displayed in the resulting forest change maps.

For example, if a pixel meets the criteria of deforestation in both the 2nd time step and the 4th time step, the final map displays deforestation occurring in the 2nd time step.

**Most recent change:** The most recent detected events of deforestation and disturbance within the time series are displayed in the resulting forest change maps.

For example, if a pixel meets the criteria of deforestation in both the 2nd time step and the 4th time step, the final map displays deforestation occurring in the 4th time step.

The resulting deforestation and disturbance maps are saved with _deforestation and _disturbance appended, respectively, to the base filename you define.

**Batch processing**

The deforestation and disturbance step supports batch processing, which enables you to process as many as 1000 images at a time. In contrast to calculating deforestation and disturbance on 2-10 input files, batch processing generates deforestation and disturbance results for each sequence of images. It also enables users to calculate deforestation and disturbance on many sets of images, in various locations simultaneously. Therefore you are not limited to processing a specific geographical location. Each line of inputs in the batch file produces a deforestation and disturbance image. This permits you to calculate maps based on different artifact removal thresholds for both deforestation and disturbance for the same geographical area, if you are unsure which threshold would be the best choice for that specific geographical location. The batch processing is initialized using a text file that contains 8 fields; each field should contain the following [from left to right]:

Fractional cover file year 1, path name and file name [input]
Fractional cover file year 2, path name and file name [input]
Reflectance file year 1, path name and file name [input]
Reflectance file year 2, path name and file name [input]
Percentage of deforestation artifact removal
Percentage of disturbance artifact removal
Deforestation and disturbance image, path name and file name [output]
Aggregation feature (on-“1” or off-“0”)
GEOTIFF file (yes-“1” or no-“0”)

Note that batch processing is also supported for non-Landsat sensors, but requires a text file with only the following fields:

Fractional cover file year 1, path name and file name [input]
Fractional cover file year 2, path name and file name [input]
Deforestation and disturbance image, path name and file name [output]
Aggregation feature (on-“1” or off-“0”)
GEOTIFF file (yes-“1” or no-“0”)

**OUTPUT**

The output of Step 4 in CLASlite is a pair of classified maps representing deforestation and forest disturbance. If more than two images are used, the output files represent deforestation for all intervals examined over the time series.

A legend file is saved alongside the _deforestation and _disturbance files, indicating the interval during which each change event was detected. Fig. 6 displays an example of this legend file for a 3-image analysis.

0 – No change detected
1 – Change from image1.tif to image2.tif
2 – Change from image2.tif to image3.tif

---

![Image 3](image3.png)

---

**Figure 6: Forest change detection**
POTENTIAL & LIMITS OF CLASlite

The power of CLASlite rests in its unique ability to convert seemingly green “carpets” of dense tropical forest cover found in the basic satellite images into highly detailed maps that can be readily searched for deforestation, logging and other forest disturbance events.

CLASlite User Interaction

CLASlite does not provide a final “map”, but a set of ecologically meaningful images that accurately identify the amount of forest cover, deforestation and disturbance. Although CLASlite is a highly automated process, you need to become familiar with CLASlite output images. All CLASlite output images – from reflectance to fractional cover to forest change images – can be readily incorporated into digital maps via Geographic Information Systems (GIS) and other common mapping and spatial analysis software packages.

Interpretation of CLASlite outputs

CLASlite detects deforestation and forest disturbance as a change in fractional cover of PV, NPV, and S from one point in time to the next. This is a detection of physical change in forest structure and thus, the results do not explicitly indicate the cause of change. As a result, detected forest change can include both natural change (e.g. tree falls) and anthropogenic change (e.g. land conversion for agriculture). However, combined with local knowledge of the study area, maps of land cover change created with CLASlite can be used to understand spatial patterns of land use change. Both deforestation and secondary forest regrowth can be tracked by the CLASlite user. Deforestation is clearly shown as a loss of forest cover, producing bare substrate and NPV. Regrowth can be tracked by a careful account of forest recovery following clearing, which must be previously mapped.

CLASlite is not a tool for direct biodiversity monitoring. It can assist in reaching conclusions regarding biodiversity from forest presence or absence and disturbance, but it has not been designed for the purpose of direct biodiversity (species) monitoring.

Field verification of CLASlite output image is highly recommended if forest monitoring is conducted as basis for on-the-ground project development and execution.

Single- versus multi-image analysis

The capability for detecting disturbance or deforestation from a single image should be used with caution. Disturbance and deforestation are based on a change in condition from one time period to another. Although patterns of deforestation or disturbance can be inferred from a single image, a human interpretation of the results is necessary. Therefore anything detected as disturbance or deforestation from a single image should be used as a guide to further investigation and validation. Using a pair of images to detect disturbance and deforestation is better than relying on a single image.

In detecting forest change, the multi-image analysis should use images from the same time of the years considered, preferably from the same month. Otherwise changes in forest phenology might affect CLASlite’s capability for forest change detection.
Basis of decision trees

CLASlite’s decision trees are calculated based on investigative research on hundreds of images throughout the Amazon basin and Andean mountains. Therefore the decision tree thresholds may not be perfectly suitable for all regions on Earth.
USEFUL REFERENCES

SCIENTIFIC BASIS DOCUMENTS


EXAMPLE APPLICATIONS


APPENDIX I: NON-DISCLOSURE AGREEMENT

CLAS and CLASlite

The Carnegie Institution for Science ("Carnegie") is a private organization that has developed the Carnegie Landsat Analysis System ("CLAS") and the Carnegie Landsat Analysis System-lite ("CLASlite") for forest mapping. CLAS, CLASlite, its subroutines including AutoMCU, and all derivatives are protected under U.S. Patents 8189877, 20090214084 and 20120288159-A1; International Classification G06K9/62.

It is the intention of Carnegie to work with other parties in the use of CLASlite for the purpose of conducting environmental, non-profit studies and ecosystem monitoring. It is recognized that such use may require the disclosure by Carnegie of certain information ("Proprietary Information") to any such party ("User"). Proprietary Information includes, but is not limited to, all information pertaining to the basis, background, development or composition of CLAS, CLASlite or any sub-routine. Proprietary Information also includes all source code, executable programs, sample data, manuals or other written documentation. The purpose of this non-disclosure agreement ("NDA") is to protect Carnegie’s Proprietary Information and to insure the proper handling and publication of data and information that is developed in studies involving the use of CLAS or CLASlite.

Accordingly, by using CLAS or CLASlite, the User agrees to the following:

1. Any version or derivative of the CLAS, CLASlite or any sub-routine software, and any results obtained from the use of CLAS, CLASlite or any sub-routine, are to be used for non-profit purposes only and shall not be used for commercial or for-profit purposes.

2. The CLAS, CLASlite or any sub-routine software, including all source code, executable programs, and supporting digital files and libraries, are the sole property of Carnegie, and shall not be reproduced by the User or provided to any party without written Carnegie approval. The software module embedded within CLAS and CLASlite, known as the Automated Monte Carlo Unmixing ("AutoMCU") code, remains the sole property of Dr. Gregory P. Asner ("Dr. Asner"). Any effort to use CLAS, CLASlite or AutoMCU in an unauthorized way or to reproduce any portion of the CLAS, CLASlite or AutoMCU software is prohibited, and will result in immediate termination of the User’s right to use CLAS, CLASlite or AutoMCU and the taking of any other legal or civil action deemed appropriate.

3. Carnegie and Dr. Asner make no warranty as to the quality and/or accuracy of any data obtained by the User pursuant to any CLAS or CLASlite study, or the suitability of any such data for any purpose. The User assumes all risks and liabilities in the collection, interpretation and use of any data or results obtained using the CLAS or CLASlite program. The User will indemnify Carnegie and Dr. Asner against any claims or damages resulting from the User’s use of CLAS or CLASlite as provided herein.

4. Upon Carnegie’s or Dr. Asner’s request, the User will return all copies of any Proprietary Information and CLAS/CLASlite Materials which may have been provided to, or used by, the User in connection with its performance of any CLAS or CLASlite studies.
APPENDIX II: IMAGE PREPARATION FOR CLASlite

Before processing data through CLASlite, you must prepare the data for input to the software. The data supplied to CLASlite must be in the correct format and have the correct number of bands. The expected formats and characteristics of the data vary by sensor. Common steps for preparing data from any satellite are:

1. Geo-referencing the image to a UTM projection (WGS-84) ellipsoid
2. Resampling spectral bands to the same spatial resolution (pixel size), if necessary
3. Reordering bands, if necessary
4. Saving the image to GeoTIFF or ENVI format

The data requirements for each satellite sensor are listed in this section. You can use almost any image processing package to prepare your image, though ENVI and ERDAS are popular options.

1. Landsat Thematic Mapper (TM) 4, 5, and Enhanced Thematic Mapper+ (ETM+) 7

CLASlite has the capability to process images from the Landsat 4, 5 and 7 satellites. The image data must be in two files. One file must contain the data for Landsat bands 1-5 and band 7 (ordered from lowest to highest). If you also have the thermal band for masking clouds (band 6, high gain), it must be in a separate file, and must cover the exact same area and have the same pixel size as the other file. Therefore, it may be necessary to resample the thermal band from its original resolution to the pixel size of the other bands, prior to use with CLASlite. When you resample the imagery, use the nearest neighbor resampling kernel. The pixel values in all Landsat imagery used in CLASlite must be 1-byte values and must not have any atmospheric corrections applied to them.

2. Earth Observing-1 (EO-1) Advanced Land Imager (ALI)

CLASlite can process images from the Earth Observing-1 satellite, which carries the Advanced Land Imager (ALI) sensor. Only ALI level-1G data are supported in CLASlite.

ALI does not have a thermal band. Only the 9 visible, near-infrared and shortwave-infrared bands must be contained in one file with the original 16-bit integer values for each pixel. The bands should be ordered from lowest to highest.

3. Advanced Spaceborne Thermal Emission & Reflection Radiometer (ASTER)

CLASlite can process ASTER Level-1B imagery acquired by the NASA Terra satellite. ASTER images come in different resolutions. The visible/near-infrared (VNIR) imagery has 15-meter pixels, while the shortwave-infrared (SWIR) imagery has 30-meter pixels. Since these bands come in different resolutions, it is recommended that all of the bands be resampled using the nearest neighbor kernel to the lowest (30-meter) resolution of the 9 VNIR and SWIR bands. The thermal bands of ASTER have 90-meter pixels, but they are not used in CLASlite.
One potential problem with ASTER data is that the VNIR and SWIR images are collected from two different telescopes, making it possible for the two sets of bands to be misaligned. Misalignment is more likely to occur in areas with large variations in elevation. When misalignment is a problem, the best thing to do is to geo-reference the images separately, then combine them using image-to-image registration.

CLASlite requires that the radiance conversion coefficients be applied beforehand. The image processing software may take care of this step during resampling, but if not, you will need to do so manually.

4. Système Pour l'Observation de la Terre (SPOT) 4 and 5

The CAP and DIMAP formats are supported for SPOT-4 (HRVIR) and 5 (HRG) in CLASlite. The 4 spectral bands should be organized into a single GeoTIFF or ENVI file with a 20-meter pixel size.

In order for CLASlite to convert the image to reflectance, it will need to read the LEADXX.DAT file for CAP formatted images and the .DIM or .XML for DIMAP formatted images. This file should be included on the original media on which the SPOT image was received and contains gain parameters that are needed to convert the image to radiance and reflectance.

As in the case of Landsat data, the input file should be 1-byte per pixel.

<table>
<thead>
<tr>
<th>Thermal band used?</th>
<th>Landsat TM &amp; ETM+</th>
<th>ALI</th>
<th>ASTER</th>
<th>SPOT 4 (HRVIR) and 5 (HRG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal band used?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of spectral bands used</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Band order</td>
<td>1,2,3,4,5,7</td>
<td>MS-1’, MS-1, MS-2, MS-3, MS-4, MS-4’, MS-5’, MS-5, MS-7</td>
<td>1,2,3N,4,5,6,7,8,9</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>Input data type</td>
<td>8-bit</td>
<td>16-bit short integer</td>
<td>32-bit floating point</td>
<td>8-bit</td>
</tr>
<tr>
<td>Processing level of data supported</td>
<td>n/a</td>
<td>Level 1G</td>
<td>Level 1B</td>
<td>CAP or DIPAM</td>
</tr>
</tbody>
</table>

Figure 7: Required data format for imagery from CLASlite-compatible sensors
APPENDIX III: ARTIFACT REMOVAL USING SLIDERS

CLASlite v3.0’s artifact removal sliders were designed to improve the quality of forest change detection and to enhance your flexibility to customize the software’s algorithms depending on your knowledge of the land cover in an area. The drawback is that sliders can remove real deforestation and disturbance pixels, or preserve false positives, if not used mindfully.

Here, we illustrate examples of where the sliders can work well and where they have shortcomings. There is no prescription slider setting for eliminating a given type of artifact, but the following are common outcomes of using the sliders. Note that the Aggregation Feature was turned on for these examples.

Example 1 depicts a scene that contains both real deforestation and disturbance near a river and artifacts due to clouds. The example illustrates that sliders, even when set to 100%, do not remove real deforestation and disturbance pixels. The deforestation slider removes all artifacts when set to 50%, but even at 100% some disturbance artifacts remain.

Example 2 illustrates that real deforestation and disturbance pixels in the form of grass cover on cleared land are at risk of being eliminated by slider values exceeding default settings. This problem can also arise if deforestation is kept at the default 50% value while disturbance is raised to a higher value.
Figure 8: Sliders example 1
Figure 9: Sliders example 2