OVERVIEW OF SEMI-AUTOMATED APPROACHES FOR MONITORING NATIONAL DEFORESTATION

FOREST CARBON, MARKETS AND COMMUNITIES (FCMC) PROGRAM

FEBRUARY 2015

This publication was produced for review by the United States Agency for International Development. It was prepared by Tetra Tech.

The U.S. Agency for International Development (USAID) launched the Forest Carbon, Markets and Communities (FCMC) Program to provide its missions, partner governments, and local and international stakeholders with assistance in developing and implementing REDD+ initiatives. FCMC services include analysis, evaluation, tools, and guidance for program design support; training materials; and meeting and workshop development and facilitation that support U.S. Government contributions to international REDD+ architecture.

This publication was produced for review by the United States Agency for International Development by Tetra Tech, through a Task Order under the Prosperity, Livelihoods, and Conserving Ecosystems (PLACE) Indefinite Quantity Contract Core Task Order (USAID Contract No. EPP-I-00-06-00008-00, Order Number AID-OAA-TO-11-00022).

Forest Carbon, Markets and Communities (FCMC) Program
1611 North Kent Street
Suite 805
Arlington, Virginia 22209 USA
Telephone: (703) 592-6388
Fax: (866) 795-6462

Stephen Kelleher, Chief of Party
Email: stephen.kelleher@fcmcglobal.org

Olaf Zerbock, USAID Contracting Officer's Representative
Email: ozerbock@usaid.gov

Tetra Tech
159 Bank Street, Suite 300
Burlington, Vermont 05401 USA
Telephone: (802) 495-0282
Fax: (802) 658-4247
E-Mail: international.development@tetratech.com
www.tetratech.com/intdev

Tetra Tech Contact:
Ian Deshmukh, Senior Technical Advisor/Manager
Email: ian.deshmukh@tetratech.com

Please cite this report as:
Steininger, M.K, G. Galindo, P. Potopov, C. Souza, M. Hansen, J. Hewson, and R. Das. 2015. Overview of Semi-Automated Approaches for Monitoring National Deforestation. USAID-supported Forest Carbon, Markets and Communities Program. Washington, DC, USA.

OVERVIEW OF SEMI-AUTOMATED APPROACHES FOR MONITORING NATIONAL DEFORESTATION

FOREST CARBON, MARKETS AND COMMUNITIES (FCMC) PROGRAM

FEBRUARY 2015

DISCLAIMER

The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ACCA	Automated Cloud Cover Assessment
AD	Activity Data
BU	Boston University
BUR	Biennial Update Report
CCDC	Continuous Change Detection and Classification
CCDM	Comprehensive Change Detection Method
COST	Cosine-Theta
CV	Coefficient of Variation
DT	Decision Tree
EF	Emission Factor
EM	End Member
FCMC	Forest Carbon, Markets and Communities
FFPRI	Forestry and Forest Products Research Institute
GFOI	Global Forest Observation Initiative
GHG	Greenhouse Gas
GLAD	Global Land Analysis and Discovery
GOFC-GOLD	Global Observation of Forest and Land Cover Dynamics
GV	Green Vegetation
IPCC	International Panel on Climate Change
MMU	Minimum-mapping Unit
MODIS	Moderate-Resolution Imaging Spectrometer
MRV	Measurement, Reporting and Verification
NASA	National Aeronautics and Space Agency
NBRI	Normalized Burn Ratio Index
NDFI	Normalized-Difference Fraction Index
NDVI	Normalized-Difference Fraction Index
NDWI	Normalized-Difference Water Index

NGO	Nongovernmental Organization
NIR	Near-infrared
NLCD	National Land Cover Database (US)
NPV	Non-green Vegetation or non-photosynthetically-active vegetation
RADAR	Radio Detection and Ranging
REDD+	Reduced emissions from deforestation and degradation, plus the role of conservation, sustainable forest management and enhancement of forest carbon stocks
RMSE	Root Mean Square Error
S	Soil
SMA	Spectral Mixture Analysis
UMD	University of Maryland
UNFCCC	United Nations Framework Convention on Climate Change
USAID	United States Agency for International Development
VCS	Verified Carbon Standards
VCT	Vegetation Change Tracker
VIIRS	Visible Infrared Imaging Radiometer Suite

1.0 ABSTRACT

International agreements of the United Nations Framework Convention on Climate Change seek to reduce national, forest-related, greenhouse gas emissions via activities and policies on reduced emissions from deforestation and degradation, plus the role of conservation, sustainable forest management and enhancement of forest carbon stocks (REDD+). These agreements require countries to report changes in forest cover and carbon stocks via a process called measurement, reporting and verification (MRV). The International Panel on Climate Change has communicated quality principles for national MRV: transparency, completeness, consistency, comparability and accuracy. Recent advances in satellite monitoring of forests have led to greater automation, which is expected to better enable countries to regularly report and meet these principles. This paper provides an overview of general steps taken with these methods to produce large-area estimates of deforestation, similarities and differences among methods, where automation occurs, where analyst input occurs and implications for national implementation and the quality principles.

2.0 INTRODUCTION

2.1 REDD+ MRV

Monitoring forest cover and deforestation is critical for effective LU planning in forested countries and for estimating the impacts of forest-related land-use change. It is also a requirement for reporting obligations associated with agreements on reduced emissions from deforestation and degradation, plus the role of conservation, sustainable forest management and enhancement of forest carbon stocks (REDD+). These agreements include those within the United Nations Framework Convention on Climate Change (UNFCCC), bi-lateral agreements and agreements with donor institutions.

In the context of REDD+, these monitoring systems are built to specifically conduct activities known as measurement, reporting and verification (MRV). MRV guidance is provided by the International Panel on Climate Change (IPCC), specifically the IPCC Guidelines for National Greenhouse Gas Inventories (IPCC, 2006). While the IPCC provides detailed guidance on the estimation of carbon stocks and calculations of emissions and uncertainty, it provides general criteria for estimating land-cover change.

The IPCC Guidelines (2006) list quality principles that are internationally accepted for national MRV. These are:

- *Transparency*: Sufficient and clear documentation allows those other than the inventory compilers to understand the compilation of the inventory and confirm data quality.
- *Completeness*: Reporting for all relevant activities is complete, with data gaps clearly documented.
- *Consistency*: Reporting for different inventory years, gases and categories ensures that differences in the results reflect real differences in emissions and not artifacts of differing methods.
- *Comparability*: Reporting must be comparable with GHG inventories from other countries.
- *Accuracy*: Reporting must seek to avoid either over- or under-estimation, and uncertainties must be estimated and reduced as much as is possible.

Another important factor is latency, or the time taken to generate estimates of forest change. Satellite-based analyses may take too long and consequently be out of date upon completion. UNFCCC reporting is due every four years for national communications and every two years for Biennial Update Reports (BURs). BURs are supposed to include information for the previous year, and thus national land-cover change assessments should be achievable within one year. These principles are relevant not solely to the UNFCCC or REDD+ but also to any monitoring system that will be used for national planning or international comparative assessments.

Several available documents summarize IPCC guidance and provide additional specific information on satellite monitoring of forests. These include the Global Observation of Forest and Land Cover Dynamics Sourcebook (GOFC-GOLD, 2015), the USAID-supported Forest Carbon, Markets and Communities Program REDD+ MRV Manual (FCMC, 2015), the Global

Forest Observation Initiative Methods and Guidance Documentation (GFOI, 2015), the UN-REDD National Forest Monitoring Systems report (UN-REDD, 2015) and the Forestry and Forest Products Research Institute's REDD+ Cookbook (FFPRI, 2015). These resources provide a range of perspectives and approaches to the use of satellite data to generate AD. Another useful review of semi-automated approaches to satellite monitoring is that of Hansen and Loveland (2012).

The remote sensing research community has engaged directly with national governments to improve national capacity for monitoring, yielding case studies of applications of different approaches at the national level. However, in most cases countries have not conducted demonstrations of different approaches at the national level.

This paper summarizes some of the methods for semi-automated detection of forest cover change using remote sensing that have been developed by major research organizations. These may be considered appropriate for national applications, and this report highlights major similarities and key distinctions among them. It is intended to complement Hansen and Loveland (2012), specifically in the context of national MRV systems. We provide this for the non-specialist in remote sensing who is nonetheless familiar with fundamental remote sensing concepts and terminology and who is involved in strategies for national or sub-national MRV systems.

2.2 FUNDAMENTAL DATA CONSIDERATIONS IN SATELLITE MONITORING OF FORESTS

An overall approach to monitoring forests begins with the selection of the source of satellite data. Some key considerations in selecting satellite imagery are spatial resolution, temporal frequency, archive length and completeness, data type, geometric and radiometric characteristics, cost and future data availability.

The spatial resolution of satellite images used for land-cover mapping ranges from less than one meter to one kilometer, or coarser. There is a trade-off between spatial resolution and both cost and the frequency of data acquisitions. Data with resolutions finer than approximately 30 meters are often expensive to acquire over an entire country. They also are not collected frequently; some very high-resolution sensors may only revisit particular sites annually or may require tasking, i.e. requests for specific acquisitions to be made. Thus it can be costly and logistically difficult to obtain national coverage of very high-resolution data, and archives can have major gaps in coverage, especially within a specific year. Very high-resolution satellite imagery are presently most suitable for sample-based approaches and for assisting calibration and validation of analyses of coarser data. Some examples of very high-resolution commercial satellite data are RapidEye, Quickbird, IKONOS, WorldView-2, SPOT HRV series, CBERS HRC, GeoEye-1 and -2, DMC constellation, KOMPSTAT-2 and RESOURCESAT-1.

Coarse-resolution data (with resolutions between 250 meters and 1 kilometer) are mostly collected on a daily basis and are available free of cost. They provide a dense time series and are particularly valuable for studying seasonal attributes such as vegetation deciduousness and inundation. However, their coarse resolution limits their usefulness to quantifying deforestation, since much deforestation occurs in small patches of several hectares or less. The most prominent examples of coarse data sources are the National Aeronautics and Space Agency's (NASA) Moderate-Resolution Imaging Spectrometer (MODIS) and the more recent Visible Infrared Imaging Radiometer Suite (VIIRS) (NASA, 2015a and 2015b), and France's SPOT VEGETATION and the more recent European PROBA-V (SPOT, 2015; ESA, 2015a). Because of their frequent acquisitions, these data are most useful for near-real-time alert systems. Several systems have been developed based on data acquired by these sensors, providing rapid alerts of locations of active fires and larger clearings, as well as conditions of drought and fire risk (e.g., INPE, 2015a; NASA, 2015c; CI, 2015; FAN 2015).

Medium-resolution data represent the most common sources for national forest monitoring because they offer an optimal combination of appropriate resolution, acquisition frequency, coverage and cost, as well as other technical characteristics. Examples are Landsat (NASA, 2015d), at 15 to 30 meters, and the SPOT HRV sensors, at 10 to 20 meters (ESA, 2015b). Landsat has the longest record and most thorough archive. With a revisit time of 16 days and routine archiving of all acquisitions, compositing multiple images to enable appropriate coverage is possible in all but the most extremely cloudy parts of the world. Landsat data include bands in the three most useful spectral regions for distinguishing forests from non-forest cover types: visible, near-infrared and middle-infrared. Data are also free to the public. While the SPOT satellites have similar characteristics, these data must be purchased and have a less thorough archive. Data from the Sentinel 2 series are also anticipated to support national forest monitoring needs in the coming years (ESA, 2015c). Sentinal 2A is scheduled to launch in 2015 and is expected to provide data with characteristics similar to Landsat at no cost.

Landsat data extend back to 1972, with early launches carrying the Multi-spectral Scanning System sensor. Since 1984 satellites equipped with Thematic Mapper sensors have provided increased spatial resolution and additional useful bands in the middle-infrared and thermal regions. For any given site, data are collected as frequently as every two weeks, pending cloudiness, allowing for a characterization of seasonal changes and sub-annual events. Data from the reflectance bands in the three regions provide the principal information content possible from optical data. The visible bands can indicate leaf absorption of sunlight as well as atmospheric haze and some variability in soil color. The near-infrared band is critical for estimating green-leaf cover, given very high reflectance in this band. The two middle-infrared bands are also useful for detecting leaf cover, via absorption by water in leaves; are sensitive to canopy shading that can help detect taller forests and can indicate variations in soils and soils versus non-green leaves and branches.

All of the above examples of data are passive optical data, i.e., they record reflected solar energy from the earth's surface in the visible, near-infrared and middle-infrared spectral regions. Another major type of satellite data is radio detection and ranging (RADAR). RADAR is active (the instrument sends a signal and measures the signal that returns from a surface) and works in the microwave spectrum, where energy interacts with vegetation and land very differently. RADAR penetrates clouds and thus is of particular interest in very cloudy regions. Presently, few RADAR data options exist that offer the coverage, archive, cost and technical specifications appropriate for accurate mapping of forest cover to warrant their use, except in the cloudiest regions. Radar and active remote sensing applications are in research and development, and better options will be available in the near future. Optical data currently remain the main data source for national forest monitoring; among these, Landsat is by far the most consistently used—understandable given its balance of high quality and free distribution.

Most satellite-derived maps of national forest cover or forest change to date have been produced via supervised classification on a scene-by-scene basis. These are often based on direct classification of change from two dates of imagery, rather than comparing the outputs of two classifications of individual dates. They have relied on analyst expertise in image interpretation, supplemented by information from the field or aerial surveys. Classifications are usually an iterative process until an acceptable product is obtained for a given scene, followed by combining all scenes to produce a national mosaic. Other approaches have applied unsupervised classification, which identifies clusters of data that are spectrally similar, or segmentation, which identifies groups of neighboring and spectrally-similar pixels. In these, the analyst interpretation occurs when they assign class labels to the clusters or segments rather than in the training process in supervised classification (e.g. Beuchle 2011, INPE 2015b).

While these products have often provided accurate estimates of forest cover, and to a lesser degree forest change, they are arguably too sensitive to variation in analyst interpretation, risk unreported errors in certain steps and are potentially difficult to reproduce. They also may show significant differences between scenes (which are processed independently), often indicated by scene edges apparent in the maps. Depending on the approach, they may be too slow and too costly in terms of analyst time for practical application on a yearly basis. For these reasons, approaches that involve greater automation and standardization of the classification process are sought to better meet the five IPCC criteria.

3.0 SEMI-AUTOMATED APPROACHES TO MONITORING FORESTS

Each of the approaches described below contains automated steps but also requires some analyst interaction during the process; no fully automated approach has yet produced validated estimates of deforestation over time at the national level. Some of the semi-automated approaches have been applied by national governments and have successfully produced national maps of deforestation that are available to the public for review. Others are methodologies currently being researched that have not yet generated products available to the public.

We restrict this review to Landsat-based approaches, since Landsat is the most common data source used for national-level forest monitoring to date. Landsat is a series of satellites that provide images in the visible, near- and middle-infrared spectra. They are delivered as image tiles with a width of 180 kilometers and a pixel layout with a 30-meter resolution. They are very well archived and researched and are available free to the public (NASA, 2015d).

Figure 1 summarizes the overall steps required for all of the methods discussed in this report. As will be discussed, there is potential for automation in all steps; in current applications, automation mostly occurs in all steps except for preparing classification. Overall processing streams used for forest monitoring can be considered as three broad phases: pre-processing, classification and post-processing.

3.1 IMAGE PRE-PROCESSING AND AUTOMATION

The quality of Landsat data has gradually improved over the past two decade as new satellites with improved sensors have been launched. Improvements also have been made in data processing for distribution to the public, as well as in methods available to users to pre-process data before applying classifications of land cover. Two key improvements are geometric and radiometric corrections. All new images from Landsat, as well as the entire historical archive, have been re-mapped to improve the geographic location of the images. As a result, the images received are now precisely mapped with a locational error of less than one pixel. This allows rapid overlay of images for estimating change without a need for users to co-register them.

Improvements in methods for radiometric corrections have reduced the effects of atmospheric conditions and of sun and view angles, enabling more consistent spectral analysis over time. This is another value of Landsat data, since they are radiometrically calibrated over time and parameters are provided to convert to comparable estimates of surface reflectance (e.g., NASA, 2015e). Automated cloud-masking tools developed for Landsat can be incorporated into a larger image-processing stream.

Data quality improvements have been complemented by increases in computer processing speeds. Programs can now process and classify large batches of images rather than just a few at a time. This enables both data-mining approaches that maximize the use of time series and approaches to processing and classifying mosaics of neighboring images over a region, or even an entire country.

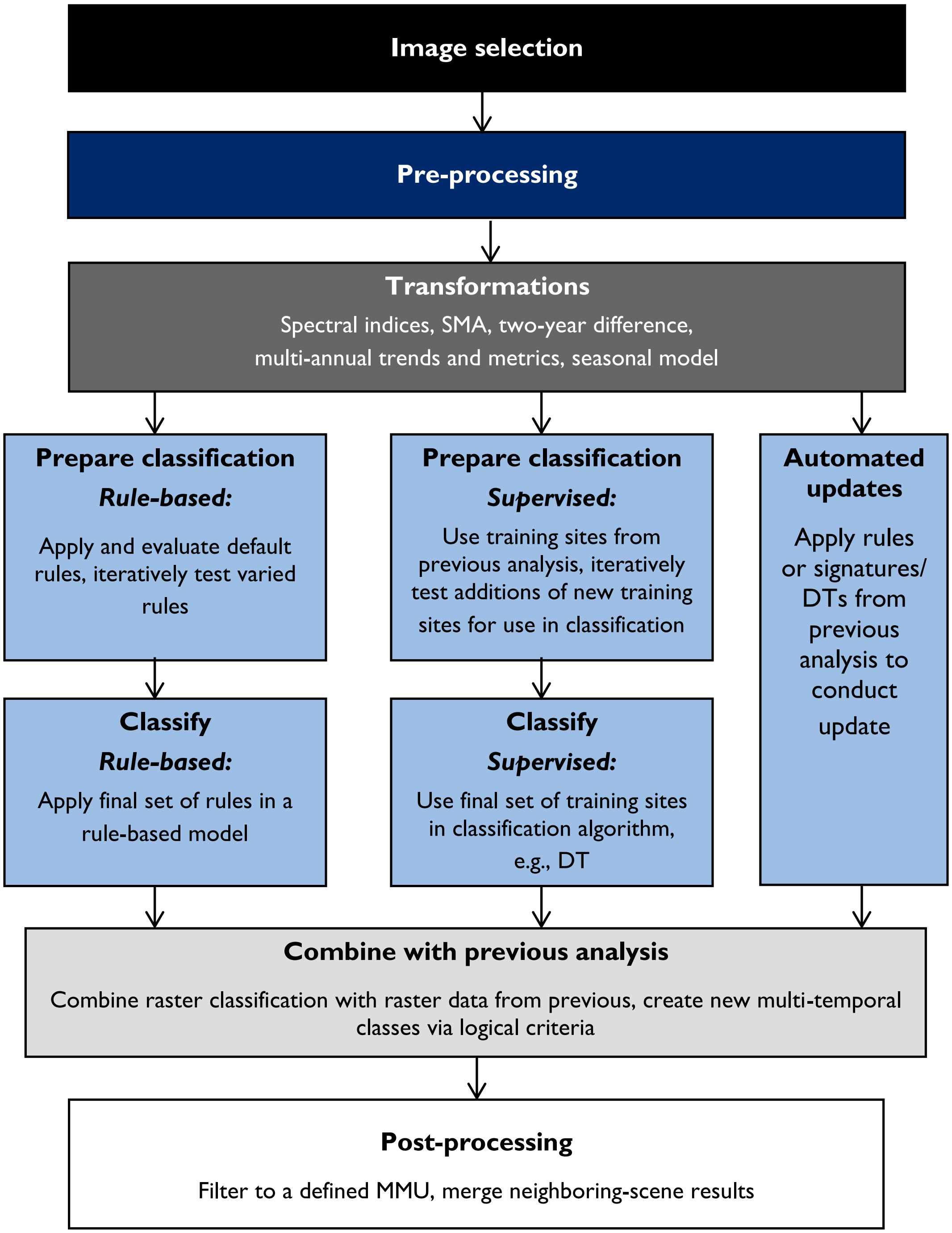

Figure 1. Flow chart of overall steps common to methods for estimating or updating deforestation or land-cover change. Currently, automation occurs in all steps except for preparing classifications. In this step, users typically interpret images to define rules or training sites for use in a rule-based or supervised classification. Interpretation must be aligned with the national definition of forest, and should be as consistent as possible with interpretation of other

areas and time periods. Although not yet applied at the national level, there is potential for full automation of updates, indicated in the box to the right.

3.2 DATA TRANSFORMATIONS AND TIME SERIES ANALYSIS

The above-noted improvements enable greater automation, at least of the pre- and post-processing steps. More automation, even if only for pre- and post-processing, means that much of the processing stream can be applied more consistently by different analysts over time. It also should result in images with fewer radiometric effects, facilitating consistent image interpretation in the classification step.

For some approaches, there is an additional step after pre-processing and classification. This is transformation of the data from reflectance to some other units (e.g., spectral indices or spectral-temporal indices). This step may further reduce the risk of inconsistent interpretation or classification outputs, as the effects of artifacts in individual bands can be reduced when combining with other bands or image dates. Consistent radiometric characteristics resulting from pre-processing, and the use of indices if data transformation is applied, can potentially facilitate semi-automation of the classification step itself.

Among proposed approaches to semi-automated monitoring—perhaps the area of greatest diversity—is in data transformations following pre-processing and preceding classification. Some of the most common spectral indices in remote sensing are vegetation indices. These are usually combinations of reflectance in the red band, highly absorbed by green leaves, and the near-infrared band (NIR), highly reflected by leaves. For example, the simple ratio of reflectances in the red and near-infrared (red/NIR) and the normalized-difference vegetation index (NDVI = NIR-red/NIR+red) are used by several methods. A significant reduction is most likely if an image is acquired immediately after clearance, when there is exposed soil, i.e., before crop development, pasture establishment or some other form of vegetation regrowth has occurred. One of the risks of being overly dependent on vegetation indices for monitoring change in forests is that many non-forest vegetation types have index values as high as those for forests, making it difficult to distinguish forest and non-forest vegetation.

Two of the approaches discussed in the Section 4 apply data transformation via spectral mixture analysis (SMA). SMA is a method to convert data from images of spectral reflectances to fractional images, with values within each pixel of sunlit leaf, soil, non-green vegetation and shade, for example. The intent is to use indices that can be commonly derived from different data sources, as well as to interpret data in terms of physical parameters that may be easier for non-specialists to understand than spectral reflectance values.

A second class of data transformation is the calculation of temporal indices. In these transformations, the differences in spectral reflectance values between two individual images from different years are calculated, either as absolute values or normalized across a range for a particular image or study area. Although difficult in cloudy regions, it is preferable to capture images from the same season when mapping change in deciduous forests. This is to minimize the potential for confusion between differences due to land-cover change versus varying drought effects between years. The use of images throughout a given year to characterize the seasonal pattern in spectral reflectances of different land-cover types also can result in a "signature" for a certain type of land cover. Fourier analysis and other methods can be used to summarize seasonal signals, and changes in derived indices over years can be estimated as a basis for estimating change in cover over time. This may involve smoothing seasonal signals or interpolation when there are data gaps for parts of the year.

A third data transformation example involves data mining and calculations of a potentially large number of temporal indices. In this computationally intensive approach, most or all images over a time period for a given area are included. Pre-processing and cloud masking are applied to maximize radiometric consistency. Images from a given year may be combined to produce mosaics that minimize gaps due to clouds. While the mosaics themselves can be classified, there is much useful information in the full set of source images, and a wide range of temporal metrics can be derived from these. Examples are change in mean reflectance among years, difference between reflectance in one year and the maximum reflectance observed in a later year, trend in reflectance over several years, etc. Particular strengths of a data mining approach are the abilities to minimize gaps from cloud cover and to maximize the potential to observe a strong signal of change soon after forest conversion. This approach may also allow classification of changes over large areas at a time, rather than on a scene-by-scene basis.

3.3 CLASSIFICATION

For the critical classification step, most methods are either supervised or based on a set of rules defined by analysts. For the former, an analyst trains the classification algorithm by exploring the image data to identify and delineate sample training sites of different classes of land-use change. The set of sites is the basis for the calculation of spectral signatures used in a classification algorithm. The remaining unidentified parts of the image are assigned to classes based on their similarity to these signatures. A range of statistical algorithms can be used, such as maximum-likelihood or decision trees (DTs; Breiman, Friedman, & Olshen, 1984), yet they are all considered supervised because of the training step.

For rule-based methods, often called knowledge-based methods, an analyst must explore the image data to determine thresholds in data values for different image indices, which are then applied to classify the data. A set of default rules may exist but users often adjust them. This is a supervised classification process, although the eventual algorithm used is much simpler, i.e., a short set of data thresholds to define classes. Because of the simple algorithm, these methods tend to benefit most from certain types of data transformations prior to classification. In both the supervised and rule-based approaches, several iterations typically are run before completing an analysis. After each iteration, results are evaluated for conspicuous errors, and training sites or rules are adjusted to address them before running a new iteration.

There is a fundamental difference between an approach based on statistical classification and one based on rules. In a statistical approach, the assumption is that spectral differences among classes can be subtle, and the derivation of precise spectral signatures is important to best estimate these. For algorithms such as maximum-likelihood, sub-classes often are created by analysts to carefully distinguish classes, while for DTs, hundreds of sub-classes are automatically created until some minimum-error criterion is met. This generally maximizes use of the information content in the image data and is adaptable to subtle variations among land-cover classes and their spectral patterns. While training data can be applied in a consistent manner over time, this is only possible if the training step is based on a similar set of training sites that remain static over time, or a similar interpretation of spectral patterns over time.

In a rule-based approach, the assumption is that the application of a small set of rules minimizes the risk of inconsistent image interpretation over time, and that these rules are sufficient to capture the subtle spectral differences between forest and other classes. Once a set of rules is determined, it is easier to demonstrate consistency over time in the method. However, there is a greater risk for erroneous classification results, especially in more difficult circumstances (e.g., in areas of seasonal vegetation, steep slopes and young secondary forests that are subtly different spectrally from mature forests). Different sets of rules may be required for different parts of a study area. Thus, in a rule-based approach, greater emphasis must be placed on demonstrating that the use of relatively few rules does not lead to unnecessarily high error levels, while in a statistical-classification approach, greater emphasis must be placed on demonstrating that the method is applied consistently over time.

When evaluating possible approaches such as these, it is important to consider the five IPCC criteria when developing and communicating methods—in this case, potential trade-offs between consistency and accuracy. It is assumed that any approach will be well documented and thus meet the criteria of transparency. Comparability with information from other countries is dependent on transparency and an assurance that the method is compatible with the national forest definition. The final principle, completeness, is dependent on an analysis that covers the entire country, or at least the managed portion of it where reporting is required. While this may appear to be a simple matter of acquiring data for the whole country, the matter is not so straightforward for particularly cloudy regions. In such situations, multiple images, each with different patterns of cloud cover, are needed for a given area and time period in order to maximize the cloud-free area for classification. In such areas, data-mining approaches that automate the seeking of cloud-free observations are of particular interest.

3.4 POST-PROCESSING

Classifications are followed by one or more post-processing procedures before a final forest cover and change map is produced. Depending on the approach, some or all of these steps may be required.

Statistical classification may involve creating many sub-classes of forest, non-forest and forest clearance. An initial step in this situation would be to merge sub-classes from the final classification into the final broader classes desired for reporting. If individual images are classified, then the results must be combined into a classification mosaic. Overlapping areas exist among neighboring Landsat images, for example, so decisions must be made about which images to prioritize in overlapping areas when producing a mosaic. Also, each image's classification, or set of threshold rules, is often independent of the others. Disagreement in the classification results should be checked for the overlapping areas, as well as throughout the study area, to limit inconsistencies and artifacts resulting from these differences.

A logical approach must be taken to combine classifications of land-cover change with maps already derived for a previous time period. Disagreements may exist between results from the new analysis and the previous product and these must be addressed. Another temporal issue is the specific acquisition dates of the images used. Cloud-free coverage may not be possible for all areas for a particular year, and thus the date of analysis may vary by one or more year. Dates will also vary within the year. It is advisable to store the source-image dates of every pixel included in the analysis. This allows a calculation of rates that corrects for varying image dates.

3.5 NATIONAL DEFORESTATION MAPS

As part of national preparations for REDD+, countries are developing their capacity to produce precise, accurate maps of forest change at the national level. Several countries have produced maps using more traditional approaches that rely heavily on analyst interpretation of individual images. These have yielded validated products that are accurate, but that could risk inconsistent image interpretation by analysts when updated, may be produced too slowly to ensure annual or biennial delivery of results or may be too cumbersome for regular application in a national MRV system. These approaches may potentially meet the five IPCC principles and be timely, depending on how they are applied. However, they do not make use of the potential for semi-automation, which could reduce the amount of time needed to generate outputs and help minimize the risk of inconsistent application over time.

Examples of national or near-national products are those generated by the PRODES program of the Brazilian space agency for the Brazilian Amazon (INPE, 2015b). While it does not include the Cerrado, coastal forest or other forests, the program covers a very large region and has been used for analyses of yearly deforestation for over a decade. In other countries, national governments have coordinated with academia and nongovernmental organizations (NGOs) that have produced and updated national maps of deforestation.

A few examples of national applications of the semi-automated approaches are discussed in the Section 4. Most are in the research domain or have recently undergone testing in sub-national and national applications. Increases in their use in national monitoring are expected as countries develop national MRV systems. Three of the methods discussed in the following section are distinct in that they have been applied at the national, or close-to-national level; two have yielded results that are available to the public for review. These three are the main methods discussed below, but others in the research domain with the goal of potential use in MRV systems are also noted.

4.0 EXAMPLES OF PROPOSED APPROACHES

This section summarizes three approaches that have been applied or proposed for national-level forest monitoring in MRV systems. The first two, ClasLite and ImgTools, are applied on a scene-by-scene basis, use SMA and estimate change directly from image pairs. The third, Global Land Analysis and Discovery (GLAD), is a data-mining approach, is applied at a regional level, includes the calculation of many temporal indices and assigns change to particular years based on the indices throughout the time period analyzed. These are the first three approaches listed in Table 1. This table lists other approaches that have been applied in the research domain but not yet tested for application within a national MRV system.

4.1 CLASLITE

ClasLite is provided by Stanford University and has been applied in various countries, such as Peru and Colombia (Asner, Knapp, Balaji, & Paez-Acosta, 2009; Cabrera et al., 2011; MINAM, 2011). It is based on developed, no-cost software that is available to users who complete a training exercise. ClasLite works with images from two different dates, applies SMA and then assigns pixels to change or not change based on a set of rules.

ClasLite begins with an atmospheric correction using the 6S program (Vermote et al., 2015). This estimates reflectance at the earth's surface, accounting for sun angle, sensor-view angle and standard atmospheric conditions for different regions that are applied scene-wide. Earlier versions masked clouds using thresholds in brightness temperature applied to Landsat's thermal band. The latest version uses Fmask, a cloud-masking algorithm developed by Boston University (BU; Zhu, Woodcock, & Ologsson, 2012). As most cloud-mask algorithms miss thin clouds or cloud edges, additional masking is applied later during the classification step.

ClasLite applies SMA by referring to a library of spectral end members (EMs) for dominant features on vegetated land. These dominant features are green vegetation (GV), non-green vegetation (NPV) and soil (S). GV may also be called sunlit leaf or photosynthetically-active vegetation. NPV may also be called non-photosynthetically active vegetation and includes branches, dried leaves and litter. Soil is any substrate and may vary significantly in spectra, yet some averages must be selected for the study area. EMs are spectral signatures for pure samples of each of these features. The logic of SMA is similar to that of gas spectrometry, where the proportion of each feature is estimated based on the spectral signature of the sample in question and knowledge of the pure spectra of expected components of the sample. In this case, the EM spectra are from libraries, i.e., records of field-based spectral measurements. The results are images of the fraction of cover of each of the features, which can be displayed, interpreted and classified as if they were reflectance images.

The fractional image for green vegetation is the most critical in these rules. In order to reduce artifacts causing within-image variations in green vegetation, the latest version of ClasLite conducts a final re-scaling of this image. This rescaling aligns the image with the percent-treecover values from the University of Maryland (UMD) GLAD product, discussed later in this

Table 1. Characteristics of selected semi-automated methodologies for monitoring forest change. See main text for references and sub-step descriptions.

Name	Temporal approach: 1-2 images per year, time series of annual data, data mining	Spatial approach: Scene-by-scene vs. mosaic	Preprocessing: Atmospheric correction	Preprocessing: Cloud mask	Data transformations	Classification algorithm	Countries, regions produced
ClasLite	Single image	Single	6S	Fmask	SMA	Threshold rules for changes in fractions	Peru, sub-national Brazil
ImgTools	Single image	Single	LEDAPS, Carlotto	SMA	SMA	Threshold rules for changes in fractions, or DT	Sub-national Brazil, Pan-Amazon lowlands
GLAD	Data mining	Mosaic	Fitted to MODIS	Proprietary	Temporal metrics of reflectances and spectral indices	DT	Indonesia, DRC, Peru, Mexico, Eastern Europe, Global
VCT	Single image	Single	None	ACCA	Spectral indices	Threshold rules for Z-scores of annual changes in indices	US
CCDM	Two images	Single	None	Proprietary	Yearly differences and vectors of spectral indices	Threshold rules for annual changes in indices	US
LandTrendr	Time series of annual data	Single	COST	Tassled-cap thresholds	Tassled Cap, temporal trend segmentation	Threshold rules for segments of temporally fitted trends	Western US
CCDC	Data mining	Single	LEDAPS	Fmask	Multi-annual, seasonal signal disaggregation	Deviation from prediction model from preceding time series trend	Massachusetts

section. ClasLite also outputs the root mean square error (RMSE) of the SMA. The fractional values, RMSE and reflectance values are then used in the classification step.

ClasLite automates the pre-processing and data transformation steps to yield fraction images for the two dates of analysis for estimating change. Following this, the classification step may or may not require some analyst interaction. A default set of rules is provided to distinguish forest and deforestation over the two dates, as well as to mask additional clouds, cloud shadow, water and wetlands. These rules are sets of thresholds in the fractional values from individual dates and differences between the two dates, as well as in RMSE and reflectance in particular bands. The default rules of the latest version first define forest in the first date and then define forest change in two steps:

Static Forest Cover:

Forest: $GV \geq 80$ and $S < 20$ (1)

Non-forest: $GV < 80$ or $S \geq 20$ (2)

Deforestation Step 1:

$GV_1 - PV_2 \geq 25$ (3a)

or $S_1 \leq 5$ and $S_2 - S_1 \geq 15$ (3b)

or $PV_2 < 80$ and $NPV_2 - NPV_1 \geq 20$ (3c)

Deforestation Step 2 (Removing False Positives):

$PV_{1,2} \geq 80$ and $NPV_{1,2} \geq 35$ and $RMSE_{1,2} \geq 6$ (4a)

or $S_2 \geq 50$ and $S_2 < 100$ and $PV_2 > 0$ (4b)

or $NPV_2 - NPV_1 < 10$ and $abs(Refl_{1b1} - Refl_{2b1}) > 300$ (4c)

and $abs(Refl_{1b4} - Refl_{2b4}) < 700$

and $abs(Refl_{1b4} - Refl_{2b4}) > 200$

where GV, NPV and S are the percent coverage of green vegetation, non-green vegetation and soil, respectively; RMSE is the root mean square error of the SMA; Refl is spectral reflectance; subscripts 1 and 2 are the first and second image dates and subscripts b1 and b4 are Landsat bands 1 (i.e., blue) and 4 (i.e., near-infrared).

In equations 1 and 2, forest is defined as high in green vegetation and low in exposed soil. Via equation 3a, a decrease in GV captures most deforestation; via 3b, an increase in soil captures additional deforestation that appeared as early secondary regrowth in the second date and via 3c, an increase in non-green vegetation captures still more deforestation. Step 2 removes false positives caused by un-masked cloud shadows and water in image 1 or 2 (4a), un-masked cloud edges in image 2 (4b) and additional un-masked cloud rings, cloud shadows and topographic shadows (4c). The rules in step 2 appear counterintuitive yet are based on testing with many images over various study areas. Following the application of rules, a three-by-three filter of roughly one-half hectare is applied to remove small artifacts of one or few pixels.

All of the above can be automated, or analyst interaction can enter in two ways. First, the above defaults can be accepted, and an additional step taken to remove more false positives. This is based on the difference over the two dates in band 1 reflectance, and users can seek an appropriate threshold via trial and error. Second, the above rules are not accepted if evaluation of the results through comparison to analyst interpretation of the reflectance images indicates poor performance. In this case, analysts would seek a different set of thresholds via trial and error. The same sequence of rules could be used, yet with modified threshold values to produce the most acceptable results. In this case, there is a significant role of analyst image interpretation in the method's application. Finally, the latest version of ClasLite includes an option to apply the SMA and rule-based classification steps to mosaics, although it is recommended that pre-processing be applied to individual scenes.

4.2 IMGTOOLS

ImgTools is provided by the Brazilian NGO IMAZON and is available at no cost to the public. It also works with images from two different dates, applies SMA and then rules to assign pixels to change or no change.

ImgTools uses the LEDAPS (Masek, 2005) combined with an algorithm from Carlotto (1999) to atmospherically correct images. Initial cloud masking is applied using cloud and shade fractions obtained from SAM (Souza et al., 2013). Further cloud masking is applied during the SMA step.

ImgTools also uses SMA to produce fraction images for GV, NPV and S, although this approach differs from ClasLite in several points in this step. First, EMs are derived from the image itself rather than from a library of field-based spectra. The tool provides an interface for exploring and selecting extremes in the multi-dimensional image data, and signatures are calculated for selected areas. These are assumed to be pure or close enough to pure for its application of SMA. Second, ImgTools calculates the fraction of shade (Sh) across each image; the shade EM is defined as 0 for all reflectance bands. While this fraction will indicate areas of cloud and topographic shadows, the main purpose of its calculation is to account for canopy shading in the following step. ImgTools also calculates the fraction of cloud (C).

A third difference is ImgTools' unique calculation of a vegetation index called the Normalized-Difference Fraction Index (NDFI). This is akin to the NDVI, which combines the fractional components of the pixel to enhance the detection of forest degradation and deforestation:

$$NDFI = (GV - (NPV + S)) / (GV + NPV + S) \quad (5)$$

where GV, NPV and S are the percent coverage of green vegetation, non-green vegetation and soil, respectively. Unlike for ClasLite, GV, NPV and S in ImagTools do not usually sum to near one. This is mainly because much of the field of view (often more than half) is canopy shade. All of the above steps are automated, and like ClasLite, the remaining steps may or may not involve analyst interpretation.

In addition to the three differences between SMA applied in ImgTools and ClasLite, there are additional differences. First, ImgTools classifies individual dates at a time and then estimates change by comparing the resulting static classifications among dates. Second, ImgTools masks clouds in the classification step, based on the cloud-fraction output of its SMA.

Two approaches have been used in the classification step in ImgTools. The first is the application of threshold rules based on a hierarchical DT defined empirically from previous studies, as in ClasLite. In this case clouds, non-forest and water are first classified and then forest is classified as either intact or degraded:

Cloud:

$$C \geq 10 \quad (6)$$

Water:

$$GV \leq 5 \text{ and } (NPV + S) < 15 \quad (7)$$

Non-forest:

$$GV \geq 85 \quad (8)$$

Forest (within Remaining Area):

$$\text{Intact: } NDFI \geq 75 \quad (9a)$$

$$\text{Degraded: } NDFI < 75 \quad (9b)$$

where C, GV, NPV and S are the percent coverage of cloud, green vegetation, non-green vegetation and soil, respectively, and NDFI is the Normalized-Difference Fraction Index in percent. As in ClasLite, analysts can interactively adjust the threshold values or add new rules using a freely provided interface.

Following classification of individual dates, changes are estimated by comparing the results from pairs of consecutive dates. This is followed by a set of rules to remove disallowed transitions. Examples are change from water to non-forest, assumed to be associated with changing water levels rather than land use, and change from forest to non-forest near cloud edges. A final step is filtering to a minimum-mapping unit (MMU) of one-quarter hectare. Earlier versions of ImgTools used DTs in a supervised classification rather than the rule-based approach, and this is still an option in the latest version.

ImgTools has been used to create maps of yearly deforestation from 2000 to 2010, with updates to 2014 expected soon, for the entire Brazilian Amazon. Results agree closely with those from the PRODES system that relies on careful analyst interaction. ImgTools has also been applied to the entire lowlands of the rest of Amazonia for 2000, 2005 and 2010 (RAISG, 2015), and tested in Indonesia and elsewhere.

4.3 GLAD

GLAD (Potapov et al., 2012, 2014a, and 2014b) is developed by UMD and has been installed in several national-government laboratories for application in national monitoring. It is very different from the above approaches in many ways. It mines the entire data archive for a study area, is applied at the level of mosaics or entire study areas, creates a large set of temporal metrics and produces a time series of forest change for all selected dates within a study period. It does not use SMA, but instead utilizes reflectance, NDVI, and NDWI and temperature from Landsat's thermal band. NDWI is the normalized-difference water index, akin to NDVI but replaces red with middle-infrared. These are then used in a supervised classification process to estimate percent tree cover and tree-cover loss, forest cover and deforestation or other types of land-cover change of interest. While much of the process is automated, analyst interpretation occurs during the supervised classification step. However, the GLAD approach analyzes data and produces results for entire regions or countries in a single step rather than for individual image tiles that must be combined to form a national product. It has yielded published, nation-wide deforestation assessments for the Democratic Republic of Congo, Indonesia and Peru (e.g., Potapov et al., 2012 and 2014; Margono, Potapov, Turubanova, Stolle, & Hansen, 2014) and a global map of tree-cover loss (Hansen et al., 2013).

GLAD begins with a calculation of reflectance above the atmosphere, i.e., not corrected for atmospheric effects but corrected for sun and view angles. It then applies atmospheric correction by normalizing the reflectance data to the atmospherically corrected data from MODIS. In doing this, GLAD fits the Landsat data to the long-term average of MODIS reflectance data. This purposely removes some of the seasonal variations in Landsat data. While using data from multiple seasons is very useful for mapping vegetation types, subtle differences within seasons may cause difficulties in estimating inter-annual changes. The approach in GLAD is to reduce seasonal signals in order to focus in inter-annual changes. This is fundamentally different from other approaches (some described in the following section) that focus on changes in seasonality from one year to another to estimate changes in cover types. GLAD has its own masking algorithm to remove cloudy and hazy pixels. As an ancillary product, GLAD outputs mosaics of cloud-free data for analyst reference in the training process.

Following pre-processing, GLAD calculates a comprehensive set of temporal metrics. This greatly reduces the data volume of an archive's time series, and users can later select which indices to use in their analyses. Metrics are in three groups and are calculated for red, near-infrared, middle-infrared, NDVI and NDWI. The first set of metrics is related to image date. Examples are the values for the mean of the first three and last three cloud-free observations, the regression slope of values versus observation date, the difference between the maximum value in the time series and previous or following minimum values and the largest change in values between consecutive dates. Second is a set of metrics related to the temporal distribution of values, representing the distributions over time as rank statistics. Examples are the values corresponding to selected ranks (minimum, 10%, 25%, etc.) and the averages for values between ranks (e.g., 10%–90% and 25%–75%). The third set of metrics is related to the values for the dates when NDVI, NDWI or temperature were greatest, least or at some rank level. Examples are reflectance corresponding to the observation when NDVI was at its minimum over the time series, or when NDWI was at a 90% rank within its temporal range. A full list is provided in Potopov et al. (2014b).

Instead of applying SMA, GLAD applies DTs to estimate either classes such as forest, non-forest and forest change, or continuous parameters, such as percent tree cover and percent tree-cover loss. DTs produce large sets of binary decisions that are optimized to minimize errors in estimation. They usually have hundreds of binary splits, creating many sub-groups of the input data that are then grouped into the final classes of interest. Referring to Section 3.3, the output of DTs serve the role of the spectral signatures that are then used to determine the most likely class for all pixels in question. DTs are similar to rules in that they are a series of splits of input data, yet they are different in two ways. They are much larger, and computers determine the optimal sequence of splits to minimize within-class variance or error. This contrasts the user-defined small sets of rules used in the rule-based approaches above.

DTs in GLAD are created via a supervised approach; analysts enter training sites for classes of interest. This is an iterative process of defining training sites, classifying data, evaluating the results and modifying training sites until an acceptable product is obtained. The algorithm first classifies whether there has been any forest change over the study period. It then assigns the year in which the change occurred based on evaluation of minimum annual NDVI throughout the period.

The application of GLAD requires initial preparation, i.e., the logistics of data acquisition, staging, pre-processing and calculation of metrics. This can be done for entire countries given sufficient computing speed and random access memory. Analyst interpretation is then required, although at a national or regional level instead of for each scene. This is comparable to analyst interpretation required in traditional supervised-classification approaches, or in the previous approaches if alternatives to default rules must be sought on a scene-by-scene basis. Once this is done, a map of forest change can be produced in one day for most countries (depending on country size and computing memory and processing speed). The resulting national classification then can be assessed and reproduced again with modified training data, as in a scene-by-scene approach. Because GLAD analyzes data summarized from the entire archive, it does not require selection of least-cloudy or most-appropriate season of images for particular years, and it automatically attributes any classified deforestation among all years of the study period.

Because of the approach to normalization of the image archive, and because the archive has been translated into a set of temporal metrics, it is possible to apply a derived DT to a new time period once the new data have been similarly pre-processed and transformed into the metrics. Or new training sites can be added to the existing set of training sites only for new areas of change. Both of these represent a high level of methodological consistency over time. GLAD can be applied to the estimation of other types of land cover and change as well, as long as analysts can provide training sites.

4.4 OTHER APPROACHES IN RESEARCH

While there is much research on automation of land-cover mapping, the examples in this section are associated with national agencies or national mapping processes. Of the four noted, three are based on changes in a set of spectral indices between two selected years, one detects yearly trends and seeks the appearance of new trends and one mines data to explore changes in seasonal signals between years. The former two are most similar to ClasLite in that they rely on spectral indices, their differences between years and thresholds or some other criteria to define change in or persistence of a land-cover type. The third uses single images from each year yet applies a trend-detection approach over a long-term time series. The fourth is most similar to GLAD in that it attempts to make the most use of the data archive and explore temporal patterns, although its temporal analysis is specific to the seasonality of different cover types and detection of changes in seasonal signals.

1. Vegetation Change Tracker (VCT). The VCT is based on changes in spectral indices (Huang et al., 2010). VCT is conducted by UMD, in support of the North American Carbon Program (NACP, 2015), but it does not involve GLAD. VCT works on individual Landsat scenes, using a single image from the summer growing season for each year of study. VCT constrains the analysis to within a baseline forest mask. It corrects the images for sun and view angle yet does not apply an atmospheric correction. Clouds are masked using the Automated Cloud Cover Assessment tool (ACCA) created for the Landsat program (Irish, 2000; Irish, Barker, Goward, & Arvidson, 2006), and additional water masking is conducted. It does not use SMA or many spectral indices as in the above methods but instead analyzes the six reflective Landsat bands and the Normalized Burn Ratio Index (NBRI), an additional spectral index based on the near-infrared and middle-infrared bands. The main algorithm produces an index of spectral difference for each pixel that is relative to each scene's forested area. This is done by calculating the difference in a pixel's reflectance from the mean for the scene's forest for each band, dividing by the standard deviation for the scene's forest and integrating the results from all bands. This results in an image of Z-scores, i.e., a relative measure of how close each pixel is to the average for forest pixels.

While there is no atmospheric correction, the assumption is that constraining to the summer growing season and calculating relative indices normalizes the data to make them comparable over time and space. The consistency in the Z-scores for a site can be compared over time, and if a pixel's Z-score becomes high (greater than three) and remains there for two years, it is assigned to the forest disturbance class. In other words, if a site becomes spectrally very different from the average for forest and remains so for two years, disturbance is assumed to have occurred.

2. Comprehensive Change Detection Method (CCDM). CCDM is also based on changes in spectral indices (Jin et al., 2011). It is used in support of the US National Land Cover Database (NLCD, 2015). CCDM's purpose, quite different from the rest of the methods discussed here, is to identify the maximum potential area of change rather than to seek the best estimate of land-use change. It is designed to be used together with the NLCD land-cover maps of 2006 and 2011 to estimate where changes have most likely occurred.

This method uses two images from each year, one from the leaf-off season and one from the leaf-on season. Spectral changes detected are assumed to indicate sites of biomass increase, decrease or stability. As in VCT, angle effects are corrected and atmospheric effects are not. CCDM applies two algorithms to two years of data separately for each season. The first is the Multi-Index Integrated Change Analysis (MIICA) that is based on changes in the values of four spectral indices. Two are the NDVI and the NBRI, noted above. The third is the Change Vector, i.e., the sum square of reflectance changes in all bands. The fourth is the Relative Change Vector Maximum (RVCMAX), which is the coefficient of variation (CV) for a pixel's changes between years in all bands divided by the maximum CV. The second algorithm divides the scene into four zones, those with high versus low values of NDVI and NBRI and increases versus increases. The results from MIICA for both seasons are then ranked and combined with the zones to estimate areas of most probably change, i.e., increase or decrease in biomass.

3. LandTrendr. LandTrendr (Kennedy et al., 2010; Cohen et al., 2010) was developed by Oregon State University and the US Forest Service. This approach takes individual images from each year, restricted to the mid-summer (middle July to late August) as much as possible. It applies the Cosine-Theta (COST) atmospheric-correction algorithm (Chavez, 1996) to an initial image, and compares dark objects over time to adjust later images to match to the base image. It uses a traditional set of spectral indices called Tassled Cap indices, which are published, consistent transformations derived from principal component analyses of many Landsat images (Crist, 1985). Indices are brightness, greenness and wetness and usually account for over 95 percent of the information content in Landsat images. LandTrendr also uses the NDVI and NBRI, as in several methods described above.

Following calculation of indices for each year's image, linear trend lines are fitted to multi-year data, the example from Kennedy, Yang and Cohen (2010) being 20 years. It then applies linear models to the time series of each index to determine temporal trajectories. Multiple models are produced, ranging from a straight line through the entire time series to lines with up to four linear segments. A statistical test is used on the models to select an optimal model. A set of rules are then applied to characterize the likelihood and type of estimated changes. The approach requires a long time series to determine trends, yet after that the model can theoretically can be updated yearly via automation.

4. Continuous Change Detection and Classification (CCDC). CCDC, developed by BU, makes use of the entire Landsat archive within a study area and time period, as in GLAD. It uses all of Landsat's reflective bands rather than derived indices. Data are atmospherically corrected using the LEDAPS tool and cloud-masked using BU's Fmask tool, as in ClasLite. Further masking of cloud, cloud shadow and snow is applied by detecting outliers during the time-series analysis step.

The model requires 15 temporal observations for any given pixel to determine a trend. Temporal data are smoothed as part of the trend analysis. Trends are disaggregated into seasonal, gradual multi-annual and abrupt change to a new trend. A statistical test is performed to determine an abrupt change, and this must be confirmed via three consecutive observations. At this point change is assigned and a new trend is estimated for the site. As with GLAD, the method is supervised, with analysts providing training sites that can be modified iteratively. It is also similar in its ability to classify multiple types of change, if training sites are provided. CCDC also requires an initial investment in accessing the data archive and setting up the pre-processing steps.

5.0 DISCUSSION

5.1 RANGE OF APPROACHES

The methodologies described in this report share main steps, although how each one is conducted varies. Most approaches calculate a set of spectral indices, which may be derived from individual images, pairs of images from two dates of analysis or time series of data. These approaches then are either supervised, where analysts enter training sites for different classes, or rule based, where analysts may still be required to interpret images to define most appropriate thresholds in indices to define rules. On one end of the range is image-by-image, rule-based approaches that require selection of images with few clouds and combining the results of neighboring classifications to produce a national product. On the other end, data-mining approaches require accessing a large archive of data and then staging and pre-processing the data, followed by analysis of an entire time series and/or region at once.

All methodologies apply corrections for sun and view angles and almost all apply atmospheric correction. Most are based on freely available tools. GLAD matches Landsat data to MODIS data, which are already atmospherically corrected. VCT uses indices that avoid the need for atmospheric correction. All methods use some form of automated cloud masking, either with a freely available tool or a simple, index-based rule, and some apply additional masking of clouds in the classification step.

Following pre-processing, methods vary in the calculation of indices used for inputs to classification. Some are ratios of reflectances in different spectral bands, SMA applied to individual dates and two-date differences in reflectances or spectral indices. Others are multi-temporal indices that are focused on seasonal signals or multi-annual trends.

For the classification step, ClasLite and ImgTools are both rule-based, i.e., they define classes via a set of thresholds applied to the indices output from SMA reflectance values. Both offer the option of iteratively adjusting the thresholds or defining new rules if the default outputs are not acceptable. ImgTools also offers the option of applying supervised DTs to the indices instead of the rules set. VCT applies a threshold to define outliers in the overall distribution of change in reflectance between dates of analysis. CCDM is based on changes between dates in a set of indices, ranking them to estimate areas of probable change. The input images for ClasLite and ImgTools are user-selected individual images from each year of analysis; for VCT they are individual images constrained to a specific season and for CCDM they are two images from each year, one from the leaf-on season and one from the leaf-off.

In contrast to the above approaches, LandTrendr, CCDC and GLAD use time-series analysis. LandTrendr constrains individual images from each year to a specific season. It then statistically determines linear trends over years and identifies breaks and the appearance of new trends to define changes in cover type. CCDC also identifies changes in trends; however, these are trends in seasonality. CCDC mines the Landsat archive to produce the densest seasonal signals possible, especially important in cloudy areas. Both LandTrendr and CCDC require multiple observations to confirm a new trend, minimizing misclassifications caused by ephemeral artifacts such as an un-masked cloud edge.

GLAD can be seen as a hybrid approach compared to the above. It mines data and conducts time-series analysis via calculation of a set of multi-temporal indices that indicate temporal trends, differences and abrupt changes. It then applies a supervised DT to map change over an entire study period. It then again uses time-series analysis, analyzing the NDVI sequence, to assign changes to specific years. GLAD can be applied to large blocks of neighboring scenes rather than to individual scenes, and it has been used to classify entire countries in a single step. The data-mining approach is important in cloudy areas, and the regional approach eliminates the need for decision making when producing mosaics of classified scenes.

5.2 IMPLICATIONS FOR MONITORING

Potential for inconsistency in interpretation. The application of these methods has almost always been semi-automated. Automation is in the pre-processing and data-transformation steps, such as cloud masking and calculation of indices. Thus far, these methods have required some level of user interaction in the classification step, and this is where inconsistencies may occur, since users may vary in interpreting images.

Supervised classifications require user input to define training areas, evaluate results and revise training sites iteratively until an acceptable product is obtained. Rule-based classifications require user input to adjust thresholds iteratively, and thus have a dynamic of evaluation and re-classification similar to the supervised approaches. Trend-detection approaches are in the research phase or have only been applied to the United States. In most applications, they can be expected also to require user input to adjust the criteria determining the appearance of a new trend that defines land-cover change.

Multiple analysts working in a laboratory may interpret different images differently within a particular time period of analysis. The same team or new analysts may interpret images differently again when a new time period is analyzed. To minimize these risks, care must be taken to standardize how images are displayed on screens, i.e., the band combinations and stretching of the data upon display. It is also critical to understand how different classes may appear when images are displayed differently. The team should come to agreement on interpretation, especially for difficult areas such as mountains, deciduous vegetation and secondary forest fallows. Analysts should refer to the images and classifications from neighboring scenes and from the previous analyses to ensure consistent interpretation.

Image selection and dates. Another area where inconsistency can occur is in decisions on image selection. For approaches that do not use data mining, i.e. all but GLAD and CCDC in this review, the particular images for each year must be selected. The spectral patterns of areas of deciduous forest can vary significantly even within a season. For example, rule-based approaches such as ClasLite and ImgTools could over- or underestimate change if leaf-growth or fall had begun at the time of only one of the two image dates in question and if rules are not adjusted accordingly. For cloudy areas, multiple images within a year are needed to obtain acceptable coverage. Options may be limited and it may be necessary to use images from different seasons, potentially leading to the problem noted above. How many additional images are to be acquired must also be decided; selecting more means increased coverage and increased analysis time and cost.

For scene-by-scene approaches, i.e. all but GLAD in this review, completed classifications of individual scenes must be combined to produce a national mosaic. Approximately 20 percent of neighboring Landsat scenes overlap along the eastern and western edges. Classification results over the non-cloudy portions of these areas will differ. Different criteria may be used, such as the date of the image or an evaluation of which classification is most correct. There may be dozens of tiles to combine for larger countries, and the process may be complicated and should be well documented.

Finally, source images vary by date, with implications for the calculation of rates. Images will often be far from December 31. The key is whether images are acquired before, during or after the typical forest-clearing season. In the tropics the middle to late dry season is when the least-cloudy images are available. Variation in particular image dates may cause considerable differences in estimates of inter-annual rates of clearing. In very cloudy areas this is an issue even for data-mining approaches, as there may be no cloud-free data for the latter part of a given year for portions of a study area. In any approach it is important to record the date of all source data at the pixel level in order to allow corrections for image date in any calculations of rates.

Forest benchmark versus change estimation. Given the above concerns, a key point should be kept in mind. It is advisable to separate the creation of a forest benchmark map from measurement of change over time within that defined forest area. This paper focuses on the latter. The former should be conducted first and thoroughly validated and evaluated by stakeholders before applying the latter. The benchmark should be as complete as possible. It may include a classification of forest types or may be combined with other data to delineate forest sub-types. Significant areas of secondary forest fallows or plantations should be separated from mature forest. Finally, the delineation of forest must be aligned with the national definition in order to enable comparability with other countries. Completing a benchmark allows the subsequent monitoring to focus solely on losses within the benchmark area. This is the process which must be consistently applied over time, whereas the benchmark need only be produced once.

Once a forest benchmark map is finalized, it need not be repeated. What requires consistency over time is the estimation of change within forests. In most cases, forest clearing over time is relatively easy to interpret since it is usually an abrupt change that produces a strong signal of spectral change. Exceptions may occur. An example is when very open or very deciduous forests appeared similar to the post-clearing land cover, in which the season of selected images is important. Another example is when the post-clearing land cover is already in a secondary-forest state, although this is unlikely if monitoring is conducted at least bi-annually, as is required to produce national BURs.

On the other hand, differences in interpretation are more likely for natural gradations of vegetation that are close to the forest-definition criteria boundary, such as open or short-statured woodlands, tree-covered wetlands and older secondary forest fallows. These are concerns mostly for the creation of the forest benchmark map. For these reasons, variations in analyst interpretation for monitoring clearing within a defined forest area should be manageable, and further advances towards full automation to further increase consistency is expected.

All of these methods have the potential for full automation with further development and calibration for forest changes in a given country or sub-national region. For rule-based and trend approaches, a country could potentially be stratified into different regions where the spectral patterns for different types of change are relatively consistent. These may be, for example, strata for mountains, seasonally inundated forest, different levels of deciduousness and woodland density, etc. It is theoretically possible for stable rules to be developed within each region, although this is an active area of research. The default rules in ClasLite and ImgTools are examples, although in practice they are often replaced by user-defined rules. DTs have already been demonstrated in the production of validated national estimates at once, an example of consistent methodological application over a large area. DTs could be developed for one time period and directly applied to a new one with no need for new training data. This has been tested for the GLAD methodology but not yet used to generate a published national update. Both the rule-based and supervised approaches already can be applied with some consistency by minimizing variations in rules applied over time or by carrying forward much of the training data from a previous time period to the analysis of a new time period. In these cases, detailed documentation could demonstrate the level of consistency.

Finally, we expect that within five years NASA's Landsat program will provide seasonal Landsat mosaics. This would reduce, although not eliminate, many of complications associated with cloud-gap filling, combining neighboring-scene classifications and image-date selection. However, there would still be sub-seasonal variations in reflectance patterns to address and the need to record source image data for rate estimation.

6.0 CONCLUSION

The issues discussed in this report have direct relevance to the IPCC principles listed in the introduction.

Transparency: It is important that methodological details for each a step in the various approaches to satellite monitoring of forest change are clear to any external reviewer, available to the public and sufficient for a new team to replicate or update the analyses. Given the above discussion, this should include provision of ancillary data, along with documentation and final product, to include dates of the source imagery used, how completed classifications were combined into a national product and how data gaps and varying image dates were addressed when calculating rates for different strata. While not the topic of this paper, the particular methods for estimating errors and any adjustments for bias should be reported. Each country team should provide documentation of the rules set and how rule-based approaches may have varied over the country, and training site files for supervised approaches. Additional helpful information would be graphic examples of the implementation of the rules or the training site maps along with the underlying interpreted spectral images. A reviewer who can see this across a country and reporting periods can better assess consistency.

Completeness: The main issue for this principle is filling gaps, mostly caused by cloud cover, to analyze the entire country or a managed portion of the country. Some areas may be consistently cloudy and may require other data sources or field surveys if considerable forest change is believed to occur there. In scene-by-scene approaches, the minor additional reduction in error in estimation of rates does not warrant the cost of additional images to obtain greater coverage.

Comparability: This is mainly a matter of aligning the analysis with the national forest definition, mostly addressed when creating the forest benchmark map. However, providing samples of rules applied or training sites along with source imagery can help demonstrate comparability. Additional high-resolution commercial data, where individual trees can be observed, or aerial or field surveys also may help in demonstrating alignment.

Accuracy: This is a matter of comparison with reference data, which may come from high-resolution imagery, or aerial or ground surveys. A statistically robust sampling scheme is important, and one should consider effects of bias in the validation process.

Latency: While not one of the IPCC principles, adequate latency is a requirement if countries are to report the most recent changes on a bi-annual basis. It is here that full automation will be most important. If full automation is not possible, approaches that involve regional or national, bulk processing, such as country-wide classifications, can facilitate timely reporting.

Consistency: This principle is perhaps the most important to demonstrate, as inconsistencies raise concerns that reported trends over time are in part due to methodological variations. The various methods discussed in this report all have shown improved consistency over time through automated pre-processing and data-transformation steps. They also seek to improve consistency by standardizing rules or enabling training sites or resulting DTs to be applied to new time periods of analysis. In current applications, all methods have had some level of analyst interaction, and thus the above discussion of inconsistency in interpretation is most relevant to this principle.

Consistency in the MRV context remains vaguely defined. For some it may mean very strong constraints on methodological variations, which is unrealistic in an evolving technical field. For others it may mean demonstrating similarities among reporting dates in parameters such as forest definition, MMU, error and bias, regardless of methodological variations. Further discussion within the MRV community on a definition of and criteria for demonstrating a consistent, national forest-monitoring system is warranted.

7.0 LITERATURE CITED

Asner, G. P., Knapp, D. E., Balaji A., & Paez-Acosta, G. (2009). Automated mapping of tropical deforestation and forest degradation: CLASlite. *Journal of Applied Remote Sensing,* 3:033543.

Beuchle, R., Eva, H. D., de Miranda, E. E., Holler, W. A., Oshiro, O. T., Achard, F., & Passarinho, A. S. (2011). Global tropical forest cover change assessment with medium spatial satellite imagery using a systematic sample grid-data, methods and first results. In *Embrapa Monitoramento por Satélite-Artigo em anais de congresso*, in *Simposio Brasiliero de Sensoriamiento Remoto* 15 Curitiba, São José dos Campos: INPE, 2011.

Breiman, L., Friedman, J. H., & Olshen, R. A. (1984). *Classification and regression trees.* Wadsworth.

Cabrera E., Vargas D. M., Galindo G., García, M. C., Ordoñez, M. F., Vergara, L. K., Pacheco, A. M., Rubiano, J. C., & Giraldo, P. (2011) *Memoria tecnica de la cuantificacion de la deforestacion historica nacional - Escalas gruesa y fina*. Bogotá, Colombia*: Instituto de Hidrologia, Meteorologia, y Estudios Ambientales*. ISBN 978-958-8067-46-9.

Carlotto, M. J. (1999). Reducing the effects of space-varying, wavelength-dependent scattering in multispectral imagery. *International Journal of Remote Sensing,* 20(17), 3333-3344.

Chavez, P. S., Jr. (1996). Image-based atmospheric corrections — revisited and improved. *Photogrammetric Engineering and Remote Sensing*, 62, 1025−1036.

CI. (2015). Conservation International (CI) Fire Risk System. Retrieved from http://firerisk.conservation.org/index.php

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync—Tools for calibration and validation. *Remote Sensing of Environment,* 114(12), 2911-2924.

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17, 301−306.

ESA. (2015a). European Space Agency (ESA) PROBA-V. Retrieved from http://proba-v.vgt.vito.be/

ESA. (2015b). European Space Agency (ESA) SPOT. Retrieved from https://earth.esa.int/web/guest/-/spot-hrv-ir-4074

ESA. (2015b). European Space Agency (ESA) Sentinel. Retrieved from http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2

FAN. (2015). *Fundacion amigos de la naturaleza* (FAN) Bolivia, Fire Risk System. Retrieved from http://incendios.fan-bo.org/Satrif-Bolivia/index.html

FCMC. (2015) Forest Carbon, Markets and Communities (FCMC) measurement, reporting and verification manual. Retrieved from http://fcmcglobal.org/mrvmanual.html

FFPRI. (2015). Forestry and Forest Products Research Institute's (FFPRI) REDD+ cookbook. Retrieved from http://www.ffpri.affrc.go.jp/redd-rdc/en/reference/cookbook.html

GFOI. (2015). Global Forest Observation Initiative (GFOI) methods and guidance documentation. Retrieved from http://www.gfoi.org/methods-guidance-documentation

GOFC-GOLD. (2015). Global Observation of Forest Cover/Land Dynamics (GOFC-GOLD) measurement, reporting and verification sourcebook. Retrieved from http://www.gofcgold.wur.nl/redd/sourcebook/GOFC-GOLD_Sourcebook.pdf

Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66-74.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160), 850-853.

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114(1), 183-198.

INPE. (2015a). Instituto Nacional de Pesquisas Espaciais, Brazil National Space Agency DETER system. Retrieved from http://www.obt.inpe.br/deter/

INPE. (2015b). Instituto Nacional de Pesquisas Espaciais, Brazil National Space Agency PRODES system. Retrieved from http://www.inpe.br/ingles/index.php

IPCC. (2006). International Panel on Climate Change (IPCC) guidelines for national greenhouse gas inventories. Prepared by the National Greenhouse Gas Inventories Programme. Eggleston H.S., Buendia L., Miwa K., Ngara T., & Tanabe K. (eds.). Japan: IGES.

Irish, R. R. (2000). Landsat 7 automatic cloud cover assessment. In *AeroSense,* 2000, 348-355. International Society for Optics and Photonics.

Irish, R. R., Barker, J. L., Goward, S. N., & Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm. *Photogrammetric Engineering & Remote Sensing*, 72(10), 1179-1188.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, 132, 159-175.

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12), 2897-2910.

Margono, B. A., Potapov, P. V., Turubanova, S., Stolle, F., & Hansen, M. C. (2014). Primary forest cover loss in Indonesia, 2000 to 2012. *Nature Climate Change*, 4, 730-735.

Masek, J. (2005). LEDAPS disturbance index: Algorithm description v. 1. *Algorithm Description for LEDAPS disturbance products*.

MINAM. (2011). *La pérdida de los bosques en el Perú,* Peru Ministry of Environment (MINAM), Programa Nacional de Conservación de Bosques para la Mitigación del Cambio Climático. Perú: *Ministerio de Agricultura, Dirección General Forestal y Fauna Silvestre*.

NACP. (2015). North American Carbon Program (NACP). Retrieved from http://nacarbon.org/nacp/

NASA. (2015a). National Aeronautics and Space Agency (NASA) Moderate Resolution Imaging Spectrometer (MODIS) data products. Retrieved from https://lpdaac.usgs.gov/products/modis_products_table/modis_overview

NASA. (2015b). National Aeronautics and Space Agency (NASA) Visible Infrared Imaging Radiometer Suite (VIIRS). Retrieved from http://npp.gsfc.nasa.gov/viirs.html

NASA. (2015c). National Aeronautics and Space Agency (NASA) Fire Information for Resource Management System (FIRMS). Retrieved from https://earthdata.nasa.gov/data/near-real-time-data/firms

NASA. (2015d). National Aeronautics and Space Agency (NASA) Landsat program. Retrieved from http://landsat.gsfc.nasa.gov/

NASA. (2015e). National Aeronautics and Space Agency (NASA) LEDAPS project. Retrieved from http://ledapsweb.nascom.nasa.gov/

NLCD. (2015). US National Land Cover Database (NLCD). Retrieved from http://www.mrlc.gov/

Potapov, P. V., Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Altstatt, A., Mane, L., & Justice, C. O. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sensing of Environment*, 122, 106-116.

Potapov, P. V., Dempewolf, J., Talero, Y., Hansen, M. C., Stehman, S. V., Vargas, C., Rojas, E. J., Castillo, D., Mendoza, E., Calderón, A., Giudice, R., Malaga, N., & Zutta, B. R. (2014a) National satellite-based humid tropical forest change assessment in Peru in support of REDD+ implementation. *Environmental Research Letters*, 9(12).

Potapov, P. V., Turubanova, S. A., Tyukavina, A., Krylov, A. M., McCarty, J. L., Radeloff, V. C., & Hansen, M. C. (2014b). Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sensing of Environment*.

RAISG. (2015). Amazonian satellite monitoring with ImgTools, *Red Amazonica de Informacion Socioambiental Georeferenciada* (RAISG). Retrieved from http://comunidadesdelperu.ibcperu.org/2012/09/12/presentacion-de-img-tools-monitoreo-satelital-de-la-deforestacion-en-la-amazonia/

Souza, Jr., C. M., Siqueira, J. V., Sales, M. H., Fonseca, A. V., Ribeiro, J. G., Numata, I., Cochrane, M. A., Barber, C. P., Roberts, D. A., & Barlow, J. (2013). Ten-year Landsat classification of deforestation and forest degradation in the Brazilian Amazon. *Remote Sensing*, 5(11), 5493-5513.

SPOT. (2015). SPOT-VEGETATION. Retrieved from http://www.spot-vegetation.com/

UN-REDD. (2015). United Nations REDD National Forest Monitoring Systems: Monitoring and measurement, reporting, and verification (M&MRV) in the context of REDD+ activities. Retrieved from http://www.un-redd.org/Newsletter38/ForestMonitoringandMRVLaunch/tabid/106350/Default.aspx

VCS. (2015). Voluntary Carbon Standards (VCS). Retrieved from www.v-c-s.org

Vermote, E. F., Kotchenova, S. Y., Roger, J. C., Tanre, D., Deuze, J. L., Herman, M., & Morcrette, J. J. (2015). 6S radiative transfer program. Retrieved from http://6s.ltdri.org/index.html

Zhu, Z., Woodcock, C. E., & Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sensing of Environment*, 122, 75-91.

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov