Variation in deforestation estimates: a comparative assessment of three remote sensing protocols for the case of Acre, Brazil

Karla da Silva Rocha [1,2]
Stephen George Perz [2]

[1] Universidade Federal do Acre - UFAC
Caixa Postal 500 - 69920-900 – Rio Branco - AC, Brasil
rochakarla@uol.com.br

[2] University of Florida - UF
P.O. Box 117330- 32611-7330- Gainesville - FL, USA
{rocha, sperz}@ufl.edu

Abstract

Many deforestation estimates have been derived from the Landsat platform in the past 30 years. More recently, estimates have also been produced from other orbital platforms, or using distinct methods of image processing and classification. As a result, there is often a diversity of estimates of LULCC available for high-profile study regions. Deforestation is increasingly seen as a metric for evaluating the effectiveness of environmental policies, and different estimates can be politicized by groups with interest in higher or lower estimates. Further, clean development mechanisms involving forest conservation have advanced toward implementation but require "accurate" estimates of forest cover with which to determine environmental service payments. Such issues are at play in many regions, notably the Brazilian Legal Amazon (BLA), where deforestation has proceeded rapidly, resulting in biodiversity loss and carbon emissions. This paper therefore compares deforestation estimates in the Amazon using data from multiple remote sensing studies. The main goal is to evaluate specific steps in remote sensing methodology to identify factors that account for differences in the resulting deforestation estimates. The comparison focuses on deforestation estimates from three different sources. The analysis shows differences in deforestation estimates; while there are many possible sources of such differences, in this analysis estimates vary primarily due to definitions of land cover classes.

Key Words: image processing, land-cover change, Amazon, processamento de imagem, mudanças de cobertura da terra, Amazônia.

1. Introduction

Technological advances in remote sensing, especially in the form of earth observing satellites, have made it easier for the scientific community to analyze the spatial extent of human impact on the environment, as well as naturally occurring environmental changes. Remote sensing enables large-scale observation of areas that would be inaccessible or otherwise difficult to access, making it applicable as a tool for monitoring Land Use and Land Cover Change (LULCC), since such changes are more difficult to quantify over large land areas using field methods of data collection. Satellite images play a very important role in the analysis of LULCC because they can cover large land areas with comparable data over time, both of which are important when studying forest changes (Dwivedi et al. 2005). Up to date remote sensing data can be obtained across a range of spatial and temporal scales at a reasonable cost. Some software and satellite images can be downloaded from the internet for free. As a result, numerous universities, research centers, governmental organizations (GO's) and non-governmental organizations (NGO's) around the world are conducting LULCC studies based on remote sensing data that can help understand LULCC dynamics.

However, there remain questions about the reliability and comparability of remote sensing data, since the characteristics of orbital platforms, processing protocols, and classification algorithms all vary, which may affect estimates of LULCC. This is problematic, since varying estimates of LULCC bear ramifications for assessments of land use, productivity and degradation, which in turn may inform policies in various sectors such as agriculture, forestry and environment. In Brazil, deforestation has become a central environmental question concerning the Amazon. Policy initiatives to reduce the rate of tropical deforestation became especially relevant in Brazil where forest loss is responsible for three-quarters of national carbon emissions, and contribute significantly to global warming (Stern 2006). It is therefore environmentally, economically and politically important to ensure clarity regarding remote sensing protocols when presenting LULCC estimates. More specifically, it is crucial to identify specific sources of differing estimates of LULCC, which may result from decisions made at various steps of satellite image processing and classification. A comparison of methodologies to estimate LULCC in a given area over a specific period of time can be very important to evaluate the consistency of data inputs for policy. This paper therefore compares deforestation estimates in the Amazon using data from multiple remote sensing studies. The goal is to evaluate specific steps in the remote sensing methodology in order to identify factors that account for differences in the resulting deforestation estimates. The comparison focuses on deforestation estimates from three sources: 1) INPE, Brazil's National Institute of Spatial Research, which is responsible for producing official deforestation estimates for the Brazilian Legal Amazon - BLA; 2) IMAZON, which produced its own deforestation estimates for Acre that became the focus of the 2007 controversy; and 3) a National Science Foundation-funded Human and Social Dynamics project at the University of Florida which produced independent estimates of LULCC in Acre and other parts of the southwestern Amazon. Acre state (Figure 1) was chosen as a study case as there are multiple estimates available for land cover over time in this area, and because it has been the focus of previous controversies over deforestation estimates

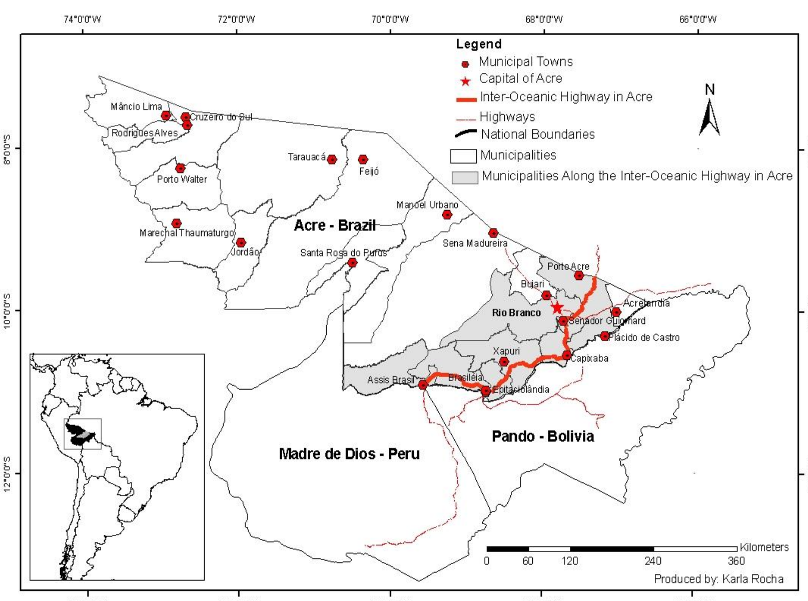

Figure 1. Study Area.

2. Methods of Deforestation Estimation in Acre from Three Sources

2.1 Remote Sensing Data Sources

Deforestation estimates from three different sources were evaluated. The INPE data come from the Brazilian Deforestation Satellite Monitoring Project -PRODES (INPE 2010). PRODES data are available for municipalities in Acre for each year from 2000 to 2010. PRODES data are also available for earlier years, but only at the level of Brazilian states. The IMAZON data come from IMAZON's 2006 report on deforestation in Acre (Souza 2006). These deforestation estimates cover a time period, from 1988 to 2004. The IMAZON study was conducted independently of INPE's PRODES program. The NSF HSD UF project data were acquired independently of INPE and IMAZON's efforts. The UF data cover a period of almost 20 years, from 1986 to 2005. However, the UF data come in 4- or 5-year time steps, unlike the INPE and IMAZON data, which come in 1-year time steps.

2.2 Image Processing Protocols

The ability to detect and quantify changes in the Earth's environment in general, and specifically of forest cover, depends on development of clear image processing protocols. Consistent protocols can help ensure accurate measurement of land cover classes and production of comparable estimates through time for accurate measurement of change. Consistency through time is specifically a challenge, as there is greater potential for similarities and differences among multiple data sets which beg questions about the processing protocols and classification methods behind a series of deforestation estimates. According to Jensen (2005), there are four fundamental steps in digital image processing of remote sensing data to extract useful information about LULCC: 1) radiometric calibration, 2) geometric correction, 3) mosaicking and 4) classification. INPE, IMAZON, and NSF HSD UF followed somewhat distinct protocols for each of these steps, making each step a potential source of differences in the resulting deforestation estimates see Table 2-2.

2.3 Land Cover Classifications

Of all the steps in satellite image processing, classification is potentially the most important for estimation of specific land cover changes like deforestation. While decisions made in the other steps may indeed result in errors and biases that can affect land cover estimation, land cover classification can have major ramifications. The classes selected, and their relationship to calculating land cover measures, can greatly affect estimates of deforestation and other types of land cover. This is especially the case insofar as there may be different land cover classes used by different sources in calculating deforestation. A key case in point concerns the definition of what constitutes forest cover, and how ambiguous classes are categorized for purposes of deforestation estimation. Secondary growth, or immature forest, may be classified as forest, non-forest, or a third category. Calculations of deforestation are affected in different classifications that separate immature forest, and if secondary growth is counted as forest or non-forest in deforestation estimation. The three sources differ in terms of bands employed in the classifications methods. The band selection is important to determine the multispectral bands optimal for discriminating one class from another. For image classification, IMAZON and INPE used bands 3, 4, 5 while the NSF HSD UF project used bands 4, 5, 7 (the near and mid-infrared bands) along with secondary derived products. Consequently, UF differs from INPE and IMAZON because UF conducted a rule-based classification instead of traditional supervised and unsupervised classifications. This technique provided flexibility to eliminate bands with striping, which limit available information from a traditional classification.

Table 2.2 Digital images processing of Remote Sensing data for three sources of deforestation estimates for Acre, Brazil: Calibration, Geometric Correction, and Mosaicking.

Basic processing steps	PRODES/INPE	IMAZON	NSF HSD/UF
1. Radiometric Calibration	Color composite images are obtained already corrected by DGI, which is in charge of receiving, processing and distributing LANDSAT and CBERS data. For radiometric calibration, DGI uses algorithms developed by Chander et al. 2009.	Algorithm developed by Carlloto (1999) implemented using EVI 4.2 software and Interactive Data Language (IDL).	Standardized method with protocol developed by CIPEC (Green et al. 1999; Green et al. 2001). Protocol implemented using ERDAS modeling and **a**lgorithm developed by Chander (2003).
2. Geometric correction	From 1997, made manually and based on official maps, which led to error propagation. Later images were registered from image to image and from 2005, using orthorectified images released by NASA	Image to image using IMAC georeferenced image year 1999 as reference image.	Image to image using University of Maryland Global Land Cover - Geocover 2000 images;
2.1- Number of Ground Control Points (GCPs) per Image	Usually 9 GCPs	35 GCPs	40 to 60 GCPs
2.2- Root Mean Square Error (RMSE)	2 pixels (90m)	Less than 1 pixel (30m)	Less than 0.5 pixel (15m)
2.3- Correction Algorithm	Polynomial Algorithm	Polynomial Algorithm	Polynomial Algorithm
2.4- Software	Spring	ENVI 4.2.	ERDAS 9.3.
3. Mosaicking	Made in SPRING, after classification	ENVI Software	Made using Erdas Software Mosaicking tool: Image Dodging, Color Balancing and Histogram Matching

INPE defined forest classes using Brazil's official technical manual of vegetation (IBGE 1992). Brazil's Institute of Geography and Statistics (IBGE) has its own vegetation classification scheme that can help classification of land cover data obtained by remote sensing. IBGE's vegetation manual distinguishes different types of forest according to type of vegetation cover. Vegetation types distinguish different forms of vegetation, such as forest trees and shrubs (Cerrado), grassy-woody (Cerrado with Clear Field), etc. This reliance on the IBGE classification differentiates INPE's classification from IMAZON and NSF HSD at UF. INPE also classified water clouds and shadows. Deforestation is generated considering previously defined classes. Where classification is based on the class attributes statistical region within certain acceptance thresholds predetermined equal to 95% or 90%, depending on the complexity of the landscape investigated (INPE 2006). IMAZON, does not remark how they classify forest cover. Beside forest they also classify clouds, shadows, degraded forest, deforestation (total and increment), beaches, sand banks, and ravine sand small formations of natural grasslands. Land-cover classes for the UF NSF project were defined in order to evaluate the impacts of road paving and other forms of infrastructure construction and upgrades on forest cover. The UF NSF project therefore considers forest and non-forest classes. The non-forest class stated in includes pasture, bare fields and urban built land cover, which were all classified as non-forest. Forest cover includes all dense vegetation cover, which includes secondary succession (generally 3-5 years of age in the study region).It is important to point out that Water, clouds and shadows classes for UF NSF project, were removed from each individual image and from the mosaics before classification. Hence while the UF classification does not include water, cloud or cloud shadows, that is because UF masked out those covers prior to classification.

The different classification schemes and definitions adopted by each source may be responsible for differences in deforestation estimates. It is therefore important to follow an established classification system instead of developing new schemes that may only be used by the producer. According to Jensen (2005), adoption of an existing broadly recognized classification system allows comparisons of the significance of classifications produced by different sources. LULC classes should therefore be selected in order to allow valid comparisons among data sources. This requires a classification system containing consistent definitions of LULC classes among sources. The classifications system in the three sources presented here is an example of a need for a standardized classification system that has to contain a consistent definition for LULC classes to permit a valid comparison of estimates.

3. Results

I highlight four points that might be especially important explanations for different deforestation estimates between sources. The first two concern the definitions of deforestation and secondary forest adopted by each source. Issues of forest and non-forest definition and classification class, constitute important factors to explain the dataset discrepancies. INPE considers deforestation to be anthropogenic modifications in mature forest for development of agriculture and cattle pasture which may give a lower deforestation estimate since only modification of mature forest is incorporated into the deforestation estimate; on the other hand, INPE consider forest regrowth or areas in process of secondary succession as deforested areas which put INPEs estimation higher when compared to other sources. Similarly, IMAZON considers secondary forest as a deforested area. IMAZON also considers as deforestation, all forest areas smaller than 0.25ha, which may be one of the reasons for IMAZON deforestation estimates to be slightly similar to INPE estimates. Different from INPE, UF NSF considers deforestation to include anthropogenic activity, all pasture areas and bare/built soil, not only areas resulting from primary forest. UF NSF also considers secondary succession as forest, since dense canopy is achieved within 3-5 years within this region. Consequently, to the extent that secondary vegetation covers eastern Acre, UF deforestation estimates may be relatively low compared to both INPE and IMAZON estimates.

The third point to consider regards classification decision concerning to clouds. INPE estimated areas deforested under clouds, while the other sources did not. Specifically, INPE assumes that the proportion of cleared areas under clouds is the same as the observable areas. By contrast, IMAZON classified clouds, shadows and water as an independent class, and made no assumption about deforestation in this class. Further, the UF NSF project removed clouds, shadows and water from each image before classification. It is possible but unclear how these differences will affect deforestation estimates by INPE as opposed to IMAZON and UF. If deforestation under clouds is higher than elsewhere, INPE estimates will underestimate deforestation; if deforestation under clouds is less than elsewhere, INPE estimates will overstate deforestation. In addition, there remain questions of the extent of cloud over, which may vary among images, even for the same path, row and year, if images come from different dates. The last point that could be the reason to expect higher or lower deforestation estimates from one source to another is the scale. Issues of scale can be relevant to explain differences between the two datasets, since coarse representation might reduce estimates of deforestation by leaving out small clearings. The coarse resolution employed by INPE, for example, leds both to underestimation of deforested areas in cases where forest clearing occurs in small plots, and to overestimation of deforestation in landscapes with small forest patches.

In summary, data presented here show that difference in deforestation estimates is associated to image processing protocols, land-cover class definition, and spatial scale, which in some cases, when associated to definitions of deforestation employed as part of the different analysis of deforestation. A clear image processing protocol can help ensure accurate measurement of land cover classes and production of comparable deforestation estimates. INPE, IMAZON and UF NSF as described in this paper, followed somewhat distinct protocols for radiometric calibration, geometric correction, mosaicking and classification, which make each step a potential source of differences in the resulting deforestation estimates. All image processing steps, like radiometric calibration, geometric correction and mosaicking may indeed result in errors and biases that can affect land-cover estimates. Differences in definitions and land-cover classification schemes adopted by each source are potentially the most important factors for estimation of specific land-cover changes like deforestation. The way that each source selected its land-cover classes, and their relationship to the calculation of land cover measures, can greatly affect estimates of deforestation and other types of land cover, this can be confirmed by analyzing the different land cover classes used by the different sources in calculating deforestation estimates.

For example, INPE classified forest cover using IBGE vegetation maps, where different types of forest are distinguished. UF NSF classified as forest all dense vegetated cover, which includes secondary secession, since dense canopy in Acre regions is achieved within 3-5 years of vegetation regrowth. On the other hand, INPE and IMAZON, different from UF-NSF, classified forest regrowth or areas in the process of secondary succession as deforested areas. The way secondary vegetation is handled by each source is a key point to answer the question of why estimates among sources are different. Definition of what constitutes forest cover, what constitutes deforestation, and how ambiguous classes are categorized for purposes of deforestation estimation is critical. Secondary growth, or immature forest, may be classified as forest, non-forest, or a third category. Calculations of deforestation are affected in different classifications as separate immature forest, and whether secondary growth is counted as forest or non-forest in deforestation estimation.

Spatial scale, when associated to definitions of deforestation employed as part of the different analysis of deforestation, may be one fact that affects estimate differences. For example, INPE covers a geographic area of 500 million ha, the entire BLA, but analyses focus on the remote sensing platform, usually Landsat, which has a spatial resolution of 30 x 30 meters, covering an area of 900 m^2. This area is afterward resampled to 60 x 60 meters. IMAZON, on the other hand, covers a small geographic area (153,149.9 km^2), the entire state of Acre. Although IMAZON uses the same remote sensing platform and spatial resolution for their deforestation analysis as INPE, deforestation definitions employed in the analysis consider areas smaller than 2,500 m^2 as deforestation, which are more than 2 pixels. This may be one of the reasons why IMAZON deforestation estimates in some cases are higher than INPE estimates.

The UF NSF project analysis covered an area of approximately 300,000 km^2, the region of Madre de Dios (Peru), Acre (Brazil) and Pando (Bolivia) - MAP region. Like INPE and IMAZON, the remote sensing data is from Landsat and spatial resolution of 30 x 30 meters, representing a pixel area of 900 m^2. Its forest definition considers all secondary successions as forest. Coarser resolution analysis therefore, can underestimate deforestation taking place in small plots, as well as overestimate deforestation in areas with remaining small patches of forest. But the issue of definition, independent of size of geographic area and spatial resolution, has decisive influence on the classification outcomes and deforestation estimates. Hence, whether a source under or overestimates deforestation is a relative issue linked to its LULC definition, in this case how they define forest, non-forest and secondary forest.

4. Conclusion

From the analysis, it is possible to say that the image processing steps followed by each source like radiometric calibration, geometric correction and mosaicking are very important to obtain data accuracy; however it is not crucial to avoid differences among deforestation estimations among sources. Differences in definitions and land-cover classification schemes adopted by each source are potentially the most important factors for estimation of specific land-cover changes like deforestation. The way that each source selected its land-cover classes and their relationships to calculating land-cover measures can greatly affect estimates of deforestation and other types of land cover. This can be confirmed by looking at the differences in land-cover classes used by the different sources in calculating deforestation estimates.

Definition of what constitutes forest cover, what constitutes deforestation and how ambiguous classes are categorized for the purpose of deforestation estimation are also crucial. INPE, IMAZON and UF NSF classify secondary growth in different ways. For example, for different sources immature forest may be classified as forest, non-forest, or a third category. Therefore calculations of what is forest and what is deforestation are affected by different classification schemes and definitions adopted by each sources resulting therefore in different estimates among sources. INPE and IMAZON separate immature forest and secondary growth from forest cover, and they consider secondary growth as deforestation; on the other hand, UF NSF includes secondary growth as forest.

How secondary growth is categorized has a huge impact on deforestation estimates. Deforestation estimates have been an important metric to evaluate the effectiveness of environmental policies, particularly programs involving payment for environmental services (PES). This raises questions by the deforestation estimates users about which available deforestation definition is best for what purpose, and if deforestation definition is a sufficient concept itself for land-cover monitoring applications, especially for PES programs.

References

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

Chander, G., Markham, B. L. & Helder, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. **Remote sensing of Environment** 113-893-903. 2009.

Dwivedi, R.S. Sreenivas K. & Ramana, K.V. Land-use/land-cover change analysis in part of Ethiopia using Landsat Thematic Mapper data. **International Journal of Remote Sensing** v. 26, n. 7, p. 1285-1287. 2005.

Green, G., Schweik, C., Hanson, M., Weber, L., Agarwal, C., Bullman, G., Carlson, L., Reisinger, M., Shields, S., Hanson, J., Barros, A.C. & Davis, B. Radiometric Calibration for Landsat 7 ETM+ scene using CPF (Equation 8) V. 1. Bloomington, IN, CIPEC: **Center for the Study of Institutions, Population, and Envronmental Change**, Indiana University.2001.

Green, G., Schweik, C., Hanson, M., Weber, L., Agarwal, C., Bullman, G., Carlson, L., Reisinger, M., Shields, S., Hanson, J., Barros, A. C.& Davis, B. Radiometric Calibration for Landsat 5 TM scenes (Equation 6) Version 1.0. Bloomington, IN, CIPEC: **Center for the Study of Institutions, Population, and Envronmental Change**, Indiana University. Ref Type: Data File. 1999.

IBGE – **Instituto Brasileiro de Geografia e Estatistica**. Manual Técnico da Vegetação Brasileira. Série Manuais Técnicos em Geociências, número 1. Rio de Janeiro. 1992.

Jensen, J.R.. **Introductory Digital Image Processing**. Prentice-Hall, Englewood Cliffs, NJ. 2005.

Souza , J. C., Verissimo, A., Costa, A. S., Reis, R. S., Balieiro, C. & Ribeiro, J**. Dinâmica do desmatamento no Estado do Acre** (1988-2004). Belém: IMAZON, 45. 2006.

Stern, N. **The economics of climate change: the Stern review**. London, UK: HM Treasury - Cabinet Office. 2006.

INPE - Instituto Nacional de Pesquisas Espaciais. **PRODES Digital**. Available at: http://www.obt.inpe.br/prodes. 2010.