AN ABSTRACT OF THE THESIS OF

Paulo Jose Murillo Sandoval for the degree of Master of Science in Sustainable Forest Management presented on March 16, 2017.

Title: Leveraging Multi-Sensor Time Series Datasets to Map Short- and Long-Term Forest Disturbances and Drivers of Change in the Colombian Andes.

Abstract approved:

______________________________________________________

Jamon Van Den Hoek

The spatial distribution of forest disturbance is commonly calculated using a satellite imagery-driven bi- or tri-temporal change analysis. Working in Colombia’s Cordillera de los Picachos National Natural Park – a region of consistent cloud cover and dramatic topographic relief – a change assessment with such infrequent observations cannot capture long-term trends of vegetative decline (browning) or improvement (greening) nor the drivers associated with these changes. In recognition of the importance of spatio-temporally explicit information for assessing the effects of socio-environmental change and conservation strategy implementation, I developed a rigorous assessment of vegetation change using MODIS and Landsat time-series data and the Breaks For Additive Season and Trend (BFAST) algorithm to identify the timing, trends, and locations of change as well the associated drivers.

First, I measured long-term vegetation trends from 2001-2015 using a Moderate Resolution Imaging Spectroradiometer (MODIS)-based 250m resolution Multi-Angle Implementation of Atmospheric Correction (MAIAC) time-series, and mapped short-term disturbances using all available Landsat images (149 dates from Landsat 5, 7, and 8). BFAST trends based on MAIAC data indicate a net greening in 6% of the park, with a net browning trend of 2.5% in the 10km-wide region surrounding the park. I also identified a 12,500 ha area within Picachos (4% of the park’s total area) that experienced a consecutive vegetative decline or browning during every year of study, a result corroborated with a BFAST Monitor assessment using finer 30m resolution Landsat data. With Landsat, I recorded 12,642 ha (±1440)
of disturbed forest within the park at high spatial and temporal accuracy. Spatially, Landsat results had user’s and producer’s accuracies of 0.95±0.02 and 0.83±0.18, respectively. Temporally, a TimeSync-supported temporal validation assessment showed that 75% of Landsat-detected dates of disturbance events were accurate within ± 6 months.

With disturbances identified, I characterized disturbances within Picachos’ southeastern foothills and associated drivers using a set of metrics related to the spectral, pattern and trend properties of disturbance patches derived from Landsat time-series data (1996-2015). A training dataset was initially developed to identify drivers of disturbances using Corine Land Cover maps and high-resolution imagery. A Random Forests classifier was used to attribute disturbances to specific drivers of forest cover change: conversion to pasture, conversion to subsistence agriculture, and non-stand replacing disturbance (i.e., thinning). Attribution of changes had high accuracy at patch and area levels with 1-5% commission and 2-14% omission errors, respectively, for regions that were converted to pasture or experienced thinning. Lower agreement was found for agricultural conversion with 43% omission and 9% commission errors.

I found that conversion to pasture is the main cause of forest cover loss within Picachos at 9901 ha (±72) corresponding to 14.7% of Picachos’ foothills, and that subtle forest alteration contributed to 1327 ha (±92) of forest degradation. Recognizing the diversity of pressures facing conservation strategy implementation in the region, these results have direct relevance for anticipating future land use pressures within Colombia, as well as across similar regions in the Andes-Amazon transition area. Indeed, since these results reveal the possibility to uncover historical disturbances related to human-incursion in protected landscapes, the methods are well suited to enhancing landscape planning particularly where biodiversity richness is quickly diminishing due to anthropogenic presence.
Leveraging Multi-Sensor Time Series Datasets to Map Short- and Long-Term Forest Disturbances and Drivers of Change in the Colombian Andes

by
Paulo Jose Murillo Sandoval

A THESIS

submitted to

Oregon State University

in partial fulfillment of the requirements for the degree of

Master of Science

Presented March 16th 2017
Commencement June 2017
ACKNOWLEDGEMENTS

I want to thank my parents Carlos and Margarita, my younger brothers Juan David and Carlos Esteban and the rest of my vast family in Colombia and other countries. All of them have contributed in many ways during this process. Thanks Cristina for your perseverance in maintaining our relationship while separated by the Atlantic Ocean, but also for listening, encouraging and always being there for me, you have been perfect.

To my friends from the beginning of this trip: Patricia, Erda and Antonio. Thanks to the Latin community that makes me feel important: Francisco, Mario and Celio; and the people who studied a lot with me: Melva, Elnaz, Tadesse and Emerson: likewise to 00’ group, my friends in Cali. I acknowledge the financial support I have received from Universidad del Tolima, which provided me with the resources to study abroad.

My special gratitude to my advisor Jamon Van Den Hoek. Thanks for teaching me and helping me during the last year. Thanks for your guidance, advice and patience, especially regarding my “amazing” English expressions. Thanks for being permanently aware of my process, and for sharing your revolutionary ideas about the linkages between our pixels and socio-political issues around the world.

This thesis would not have been possible without the vision of my initial mentor, Professor Thomas Hilker. I feel and enormous gratitude to him, for giving me the opportunity to come here, for teaching me, inspiring me and making me feel super especial for him. Regrettably, he won’t be here with us to witness the end of this stage. I really expected to make him feel proud of his Colombian student. I always will feel honored and proud for having been his friend and colleague, and I will do my best to follow his steps with the humility, charisma and scientific spirit that always characterized him.
CONTRIBUTION OF AUTHORS

Dr. Thomas Hilker and Dr. Jamon Van Den Hoek were involved in the design and interpretation of findings. Both contributed to the manuscript writing.
TABLE OF CONTENTS

Chapter 1. General introduction.....................................................................................1
  1.1 Overview .................................................................................................. 1
  1.2 Change detection methods ....................................................................... 3
  1.3 Attribution of change agent ..................................................................... 4
  1.4 Proposed Method ..................................................................................... 5
  1.5 References ................................................................................................ 6

Chapter 2. Leveraging multi-sensor time series datasets to map short- and long-term tropical forest disturbances in the Colombian Andes .................................................13
  2.1 Introduction ............................................................................................ 13
  2.2 Materials and Methods........................................................................... 15
  2.3 Results .................................................................................................... 23
  2.4 Discussion .............................................................................................. 29
  2.3 Conclusions ............................................................................................ 34
  2.4 References .............................................................................................. 35

Chapter 3. Monitoring forest disturbance and attributing drivers of change in the Colombian Andes.............................................................................................42
  3.1 Introduction ............................................................................................ 42
  3.2 Study area ............................................................................................... 44
  3.3 Methodology .......................................................................................... 45
  3.4 Results .................................................................................................... 51
  3.5 Discussion .............................................................................................. 56
  3.6 Conclusions ............................................................................................ 59
  3.7 References .............................................................................................. 60

Chapter 4. General Conclusions ..................................................................................65
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>66</td>
</tr>
<tr>
<td>Apendices</td>
<td>78</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. The study area of Picachos National Park is located in central Colombia at the northern extent of the Andes Mountain range.</td>
<td>16</td>
</tr>
<tr>
<td>Figure 2. Example Breaks For Additive Season and Trend (BFAST) output of a browning pixel. Three major breakpoints (dashed red lines) and four segments (black lines) have been identified over the 13.5-year Multi-Angle Implementation of Atmospheric Correction (MAIAC) time series (blue lines) with missing observations due to cloud effects. The slope coefficients ($\beta$) are all significant ($\alpha = 0.05$) and $P$ represents p-values.</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3. Example of pixel-level BFAST Monitor output using a 2007-2010 subset of the Landsat time series. Black solid lines denote the beginning of the monitoring period, black dots are observed NDVI observations, blue lines denote expected values from the fitted seasonal model, and grey dashed lines represent the residuals in each monitoring period. A breakpoint (dashed red line) is identified in early 2007 with a magnitude (M) of -0.35 but is not detected in 2008 or 2009. The historical periods for 2008-2009 and 2009-2010 are shorter than the 2007-2008 period because BFAST Monitor only includes preceding observations until the stable seasonal model breaks down.</td>
<td>20</td>
</tr>
<tr>
<td>Figure 4. Duration of MAIAC-based a) browning and b) greening trends based on the number of years with significant trends ($\alpha=0.05$), and (c) distribution of browning and greening periods across Picachos and its surrounding area.</td>
<td>24</td>
</tr>
<tr>
<td>Figure 5. Cumulative BFAST Monitor-derived disturbance based on Landsat 2001-2015 time series, aggregated to 1km pixel size for display.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 6. Temporal distribution of disturbance area within Picachos and the surrounding region. The year 2003 was not evaluated because of a lack of Landsat images.</td>
<td>26</td>
</tr>
<tr>
<td>Figure 7. TimeSync-based comparison of sub-annual BFAST Monitor-output disturbance dates against breakpoint dates identified using reference datasets.</td>
<td>28</td>
</tr>
<tr>
<td>Figure 8. Comparison between years of breakpoints detected using MAIAC and Landsat-based disturbances analyses in whole study area.</td>
<td>29</td>
</tr>
<tr>
<td>Figure 9. a) Disturbance pattern in southeastern Picachos; b) Change magnitude ($\delta$NDVI) subset over Platanillo sector and presence of unpaved roads (2015 RapidEye Red-Edge band background); c) Subset for Rio Platanillo, Rio Guayabero, and Rio Guaduas (2014 RapidEye Red-Edge band background).</td>
<td>31</td>
</tr>
<tr>
<td>Figure 10. Study area located over Picachos National Park foothills.</td>
<td>44</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Conceptual flowchart for disturbance detection and driver attribution</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Example BFAST Monitor-derived disturbance trajectories associated with drivers of change. Vertical red lines indicate a breakpoint, blue lines indicate a seasonal+trend model fitted and M denotes the magnitude of change.</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Most important metrics for discrimination among drivers based on mean decrease accuracy in RF model</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Relationships between time-series metrics and drivers evaluated.</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Temporal distribution of changes per class. Year 2003 is not included due to lack of Landsat imagery</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Spatial distribution of classes from 1999-2015.</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Percentage of cloud-free Landsat observation in the period 1996 –2015.</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Selection of magnitude threshold ($P(disturbances)=0.5$) using Binomial Logistic Regression. The grey area is 95% confident interval. Training sites (blue points) were selected based on visual interpretation where 0=non-disturbed and 1=disturbed.</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Area-weighted error matrix of Landsat-based BFAST Monitor results using stratified sample sites across classes: 1) Disturbance, 2) Non-disturbance outside 2km unpaved roads buffer, and 3) Non-disturbance inside 2km unpaved roads buffer.</td>
<td>27</td>
</tr>
<tr>
<td>Table 2. Summary of patch-level metrics used to characterize drivers of disturbance. SRTM = Shuttle Radar Topography Mission Digital Elevation Model (90m).</td>
<td>49</td>
</tr>
<tr>
<td>Table 3. Area-weighted error matrix of NDMI-based BFAST Monitor assessment using stratified sample sites across classes: disturbance and stable forest.</td>
<td>51</td>
</tr>
<tr>
<td>Table 4. Confusion matrix at patch level.</td>
<td>54</td>
</tr>
<tr>
<td>Table 5. Area-weighted error matrix for drivers.</td>
<td>54</td>
</tr>
<tr>
<td>Table 6. Total number of Landsat images used for disturbance detection.</td>
<td>78</td>
</tr>
</tbody>
</table>
DEDICATION

To:

Thomas Hilker
Chapter 1. General introduction

1.1 Overview

Human dominance on Earth has introduced a diversity of landscape disturbances with dramatic impacts on the environment (Achard et al., 2010). As a result, one third to one-half of the Earth’s land surface had been affected by human-induced activities (Zeng, Sui, & Wu, 2005). This landscape’s alteration is generally driven by economic development, agricultural production and demographic transitions (Leblois, Damette, & Wolfersberger, 2016). With population growth and an unbalanced economic model in both peripheral and core countries, human-induced landscape changes will continue, affecting our ecosystems. Hence, a continuous monitoring and understanding of disturbance regimes as well as the implementation of strategies to reduce human impacts are some of the biggest challenges for humankind during the 21st century (Zeng et al., 2005).

For decades, creating protected areas (PAs) has been considered one of the most effective practices to promote and ensure biodiversity conservation (Cuenca, Arriagada, & Echeverría, 2016; Wiens et al., 2009). However, global satellite studies show a dramatic increase in disturbance surrounding protected areas, specifically in tropical countries that hold most of the world’s biodiversity (DeFries et al., 2007; Defries, Hansen, Turner, Reid, & Liu, 2007; Joppa, Loarie, & Pimm, 2008). The Andes, in particular, is considered as hyper-hotspot because of its exceptional amount of endemic plants (Myers, Fonseca, Mittermeier, Fonseca, & Kent, 2000), yet methods for monitoring disturbance within such biodiverse regions are still limited (Pasquarella, Holden, Kaufman, & Woodcock, 2016).

The cumulative changes in the regions surrounding protected landscapes within the Andes are threatening the extraordinary biological diverse of many countries. For example, in Colombia, the second most biodiverse country in the world (Butler, 2006), the rapid expansion of human population and the growing demand for forest services have increased the pressure over PAs located in the Colombian Andes. Forest cover changes are related to a variety of factors including rural-urban forced migration, presence of illicit crops, legal and illegal mining exploitation, and
agriculture and ranching expansion (Dolors Armenteras, Rodríguez, Retana, & Morales, 2011; Etter, McAlpine, Wilson, Phinn, & Possingham, 2006; Sánchez-Cuervo & Aide, 2013).

While some studies suggest that PAs on Colombian Andes have been effective at limiting deforestation and degradation (Cortes Alonso & Sergio, 2012; Rodríguez, Armenteras, & Retana, 2013; Salazar Villegas, 2013; Sanchez-Cuervo & Aide, 2013), the complex colonization processes in the Andes–Amazonia transition belt (Dolors Armenteras et al., 2011) driven by internal conflict, drugs trafficking and subsistence agricultural expansion have nonetheless caused drastic modifications to forest structure and extent over the last 50 years (Dolors Armenteras et al., 2011; Etter, McAlpine, & Possingham, 2008; Sánchez-Cuervo & Aide, 2013). As a result, the Andes–Amazonia transition belt is ever more recognized as a hotspot of deforestation dominated by encroaching pastures, agriculture, and illicit crop cultivation within and outside of PAs (Achard et al., 2014; Rodríguez et al., 2013).

Even though the Landsat archive stretches back to 1972 and offers consistent measurements of the Earth’s surface over decades (Holden & Woodcock, 2015; Vogelmann, Gallant, Shi, & Zhu, 2016), most satellite image-based analyses in the Andes region use only two or three images for change detection (Dolors Armenteras et al., 2011; Etter, McAlpine, et al., 2006; Fernández et al., 2015; Rodríguez et al., 2013; Tovar, Seijmonsbergen, & Duivenvoorden, 2013). Bi- or tri-temporal comparisons are unsuited for monitoring complex, dynamic socio-environmental relationships at different spatio-temporal levels. The free availability of Landsat and Moderate Resolution Imaging Spectroradiometer (MODIS) imagery support the development of more sophisticated algorithms that extract finer-scale traits and abrupt changes using individual vegetation indices and transformation functions that take advantage of full-spectral domain (Alonzo, Van Den Hoek, & Ahmed, 2016; Hilker et al., 2012; Jamali, Jonsson, Eklundh, Ardo, & Seaquist, 2015; Kennedy, Yang, & Cohen, 2010; Rufin, Müller, Pflugmacher, & Hostert, 2015; Verbesselt, Zeileis, & Herold, 2012). Notwithstanding, we have only just begun to comprehend how to use frequent, repeatable and reproducible observations from different satellite-based sensors for study of terrestrial ecology.
In an attempt to provide a more consistent perspective of change dynamics and to overcome the obstacles for monitoring inaccessible places, this study integrates 1) multi-source imagery with a sub-annual change detection algorithm, 2) a set of diverse time-series metrics for detecting long- and short-disturbances forest, and 3) the most probable drivers of change in National Park Picachos. The methods presented here support improved comprehension and spatial, temporal, and thematic detail of the forest disturbances in regions characterized by complex socio-economic dynamics and persistent cloud cover.

1.2 Change detection methods

Forest cover change detection is the process of identifying differences in the state of forest cover condition or structure by observing the forest at different times (Singh, 1989). Early remote sensing change detection techniques were strictly bi-temporal (e.g., Coppin & Bauer, 1996; Lu, Mausel, Brondizio, & Moran, 2004; Singh, 1989), and bi- or tri-temporal comparisons continue to be used in regions with persistent cloud cover or in areas with low image availability (Alonzo et al., 2016; DeVries, Verbesselt, Kooistra, & Herold, 2015). However, such simplistic representations of forest changes are unable to capture the consequences of complex co-evolution of natural and social systems across different spatial and temporal scales (Murillo-Sandoval, Van Den Hoek, & Hilker, 2017).

In recognition of these limitations, change detection algorithms such as LandTrendr and Vegetation Change Tracker (VCT, Huang et al., 2010) have been developed that employ the full satellite data time series to identify location and timing of both abrupt and protracted changes with high accuracy (Cohen, Yang, & Kennedy, 2010; Pflugmacher, Cohen, & Kennedy, 2012; Pickell et al., 2014; Thomas et al., 2011). Annual time series detection algorithms require the reduction of potential seasonality by selecting the best-available-pixel (Hermosilla, Wulder, White, Coops, & Hobart, 2015). This temporal reduction both (over-)simplifies the characterization of forest cover change and decreases the amount of data storage and computational requirements (Forkel et al., 2013).
In an effort to extend detection of disturbances to a finer, sub-annual temporal scale, some techniques have focused on increasing temporal resolution by combining daily MODIS observations with high-resolution imagery, (e.g., Hilker et al., 2009; Tran, de Beurs, & Julian, 2016). Other approaches such as the Noise Insensitive Trajectory Algorithm (NITA; Alonzo et al., 2016), the Detecting Breakpoints and Estimating Segments in Trend (DBEST; Jamali et al., 2015) and the Breaks for Additive Seasonal and Trend (BFAST; Verbesselt, Hyndman, Newnham, & Culvenor, 2010) detect changes using all available observations at all pixels. Of these, BFAST has been broadly tested across geographies and land cover types, has simple parameterization, and less sensitivity to data gaps and noise, making it suitable for monitoring forest changes in regions with high cloud cover and low data availability (DeVries et al., 2015; Schultz, Verbesselt, Avitabile, Souza, & Herold, 2015; Watts & Laffan, 2014). Although sub-annual algorithms like BFAST require higher data volumes and computational requirements, they offer a more accurate representation of subtle forest changes and gradual vegetative trends than broader annual resolution approaches (Kennedy et al., 2014; Pasquarella et al., 2016).

1.3 Attribution of change agent

According to White & Jentsch (2001), forest disturbance is “a relatively discrete event in time that disrupts the ecosystem, community or population structure and changes the resources, substrate availability or physical environment. A disturbance, in this sense, changes the state of structural and physical variables in the ecosystem, although these changes also influence ecosystem functions and processes.” The methods described in section 1.2 are able to track the presence of forest disturbance, but equally important is to identify the agents causing disturbances. Even though the magnitude of forest cover change detected by algorithms correlates with the disturbance severity (Jeffrey G. Masek, Hayes, Joseph Hughes, Healey, & Turner, 2015), a robust assessment of agents must also include repetitive measurements in the spectral, temporal, pattern and location domains (Lambin, 1999). Methodologies for attribution of changes tend to focus on individual drivers such as pests and insects (Jones, Song, & Moody, 2015; Neigh, Bolton, Diabate, Williams, & Carvalhais,
2014; Olsson, Jönsson, & Eklundh, 2012), fires (Chu, Guo, & Takeda, 2016; Jiang, Jia, Chen, Deng, & Rao, 2015), land use conversion (Borrelli, Sandia Rondón, & Schütt, 2013; Reiche, 2015), harvest (Buhal, 2016; Hamunyela, Verbesselt, & Herold, 2016) and wind or storm damage (Al-Amin Hoque, Phinn, Roelfsema, & Childs, 2017; Nelson, Kapos, Adams, Oliveira, & Oscar, 1994) rather than multiple agents. The difficulty in discriminating multiple drivers of disturbance results from the lack of reliable, quantitative spatio-temporal datasets that support the assignment of specific changes to a specific cause.

The use of high-resolution imagery and temporally dense Landsat time-series with assistance from visual interpretation has become the most popular and reliable way to initially categorize different agents of change (Kennedy et al., 2015) yet automated methods for agent attribution are gaining ground. Pickell et al. (2014) introduce a highly accurate (94%) method using VCT and linear discrimination analysis (LDA) across a large number of Landsat-derived spectral, temporal and pattern metrics to classify wildfires and resource extraction in Alberta, Canada. A similar approach was used by Hermosilla et al. (2015) in which a time-series protocol based on the “best available pixel” was input to a Random Forest (RF) classifier to categorize disturbance drivers including fire, road expansion, forest harvest, and non-stand replacing disturbance. Disturbance attribution was also assessed by Kennedy et al. (2015) using a set of spectral and trend metrics derived from LandTrendr input to a RF classifier for discriminating agricultural conversion, urbanization, fire and insect damage, and harvests associated with forest management. These methods offer a relatively straightforward way to exploit the Landsat archive along with spectral, temporal, and pattern metrics to distinguish among different drivers of change and are improved with a longer and denser Landsat archive.

1.4 Proposed Method

The proposed method explores a different approach to the problem of sub-annual forest cover change monitoring in persistently cloudy places by i) combining multi-source remote sensing time-series data to characterize short- and long-term
disturbances, and ii) distinguishing different drivers contributing to observed changes, using machine learning classifiers in Picachos Natural Park in Colombia.

In my first chapter, I take advantage of MODIS-based MAIAC observations because of its enhanced surface reflectance and more regular temporal resolution, and, using BFAST, I classify disturbance trends as either greening or browning. Greening and browning are statistically significant linear segments between breakpoints, which exhibit a positive or negative slope direction, respectively. I also extend the study to the Landsat scale of observation using a variant of BFAST at sub-annual level, which initially employs historical observations to fit a stable model and then compares the characterization of a stable forest with other observations in order to detect significant deviations in signal values (referred to as breakpoints) from 2001 to 2015. In pairing MODIS and Landsat time series, I compare the differences in disturbance location and timing and assess the ability of both sensors to detect abrupt and gradual forest cover changes in Picachos.

In my second chapter, I collect training data at sites of forest thinning as well as forested sites that have been converted to pasture or agriculture. This training sample was collected from Corine Land Cover maps and high-resolution imagery. Every patch of disturbed forest is characterized with metrics derived from spatial pattern, spectral values, trends, or topographic condition, which are input to a RF classifier to assign the most probable driver of change to each patch. The incorporation of multi-source imagery with a sub-annual change detection algorithm and diverse metrics permits an unprecedented identification of spatio-temporal change as well as the local-level processes directly contributing to the changes.

1.5 References


Cortes Alonso, & Sergio, D. (2012). *Assessing the ground truth drivers of land cover and land use changes at a local scale in Cundinamarca and Tolima Departments in Colombia*. University of Southampton.


Thomas, N. E., Huang, C., Goward, S. N., Powell, S., Rishmawi, K., Schleeweis, K., &


Leveraging multi-sensor time series datasets to map short- and long-term tropical forest disturbances in the Colombian Andes.

Paulo J. Murillo-Sandoval, Jamon Van Den Hoek and Thomas Hilker


[http://www.mdpi.com/2072-4292/9/2/179](http://www.mdpi.com/2072-4292/9/2/179)
Chapter 2. Leveraging multi-sensor time series datasets to map short- and long-term tropical forest disturbances in the Colombian Andes

2.1 Introduction

Protected areas are vital for preserving the planet’s biodiversity and maintaining functional terrestrial and aquatic ecosystems (Wiens et al., 2009). The tropical Andes, in particular, are a “hyper“ hotspot for endemism and conservation support (Myers et al., 2000). However, protected areas in the Colombian tropical Andes are under threat from anthropogenic land uses such as cattle grazing, agriculture and population migration (Dolors Armenteras et al., 2011; Jetz, Wilcove, & Dobson, 2007; Sánchez-Cuervo & Aide, 2013), which have led to deforestation and degradation of protected landscapes (Simula & Mansur, 2011). Monitoring long-term ecological processes in these protected areas is therefore crucial to ecological conservation and biodiversity (Willis, 2015). While remote sensing presents the ideal means to frequently and regularly characterize vegetation change over large and remote areas without the need for physical access, remote sensing of the Colombian Andes is challenging due to frequent cloud cover, complex topography, and regional insecurity (e.g., armed conflict) that impede field validation. Indeed, to date, most studies on Latin American tropical forests have been confined to the Amazon basin’s lowlands rather than the Andes region (Fernández et al., 2015; Rodríguez et al., 2013).

In the Andes, as well as other mountainous regions, studies of vegetation dynamics have typically used bi- or tri-temporal satellite images acquired on relatively cloud-free dates (D. Armenteras, Gast, & Villareal, 2003; Rodríguez et al., 2013; Tovar et al., 2013). Such temporally coarse and irregular change assessments are prone to confusing seasonal (intra-annual) vegetation dynamics with inter-annual changes (DeVries et al., 2015; Van Den Hoek, Ozdogan, Burnicki, & Zhu, 2014) and are wholly dependent on the quality of images selected for analysis (Alonzo et al., 2016). In Colombia, the ecological monitoring of protected areas is managed by Colombian National Natural Parks (Parques Nacionales Naturales), and based on maps generated every five years using the Corine Land Cover inventory (CLC), which is partially or poorly validated in many remote areas (Anaya, Colditz, & Valencia, 2015; IDEAM, 2010). Such simplistic representations of vegetation change
are unable to capture the complex co-evolution of natural and social systems across different spatial and temporal scales.

In an effort to provide a more systematic perspective of vegetation dynamics and illuminate spatially detailed patterns, we employ a multi-scale time series assessment of forest disturbance over Picachos National Park and its surrounding area from 2001-2015 using two datasets: the Moderate Resolution Imaging Spectroradiometer (MODIS)--based Multiangle Implementation of Atmospheric Correction (MAIAC) time series, and the Landsat archive from 1996 – 2015 (Landsat 5, 7, and 8). MAIAC is a relatively new cloud screening and atmospheric correction technique that provides a more complete description of the physical system under investigation and overcomes many restrictions faced by conventional satellite retrievals (e.g. assumed Lambertian reflectance), especially in tropical regions (Hilker et al., 2012, 2015). MAIAC provides a more confident interpretation of vegetation trends with much greater temporal fidelity than what could be achieved with a bi or tri-temporal comparison. (Bi et al., 2016; Hilker et al., 2014), but its coarse spatial resolution does not support detection of fine-scale changes that may be better characterized using Landsat data. The Landsat archive with its 30 meter spatial resolution and 16-day revisit period, is more suitable for monitoring small-scale processes of deforestation and degradation (Achard et al., 2010, 2014) like those present in Picachos and other regions of montane in Colombia (Rodríguez et al., 2013).

To characterize time series trajectories and trends, we apply the Breaks For Additive Season and Trend (BFAST) algorithm, which has successfully been used to detect forest cover changes in tropical forests with comparable structure to Picachos (DeVries et al., 2015; Dutrieux, Verbesselt, Kooistra, & Herold, 2015). BFAST was selected over other well-known disturbance monitoring algorithms such as LandTrendr (Kennedy et al., 2010) and Vegetation Change Tracker (VCT) (Huang et al., 2010) for two main reasons. First, BFAST has been shown to be more resilient to noise and missing data (i.e., low data availability or cloud cover) (Devries, Pratihast, Verbesselt, Kooistra, & Herold, 2015; Hamunyela et al., 2016; Schultz et al., 2016). Second, BFAST produces sub-annual information on
disturbances and trends rather than the annual scale of LandTrendr and VCT (Cohen et al., 2010; Huang et al., 2010). Two different implementations of BFAST were applied: we first used the dense MAIAC time series to quantify multi-year vegetation greening and browning trends and recorded breakpoint dates, which may indicate disturbance or a significant deviation in vegetation trends (Verbesselt et al., 2012). Second, we used the BFAST Monitor method and all available Landsat imagery to identify annual disturbances (DeVries et al., 2015), which were independently validated using TimeSync (Cohen et al., 2010). Short-term Landsat disturbances were then used to validate the timing and distribution of longer-term MAIAC trends. Collectively, these techniques present an effective and scalable framework for monitoring short- and long-term vegetation disturbances and trends in the Andes and other mountainous forested regions.

2.2 Materials and Methods

2.2.1 Study Area

Picachos National Park was established in 1977 and is the third largest national park in the Colombian Andes with an area of approximately 288,000 ha. It has a pronounced elevation gradient spanning 450–3800 m, and is characterized by high annual rainfall (4000 mm/year) and a diverse mosaic of tropical ecosystems (Figure 1). Picachos also has great hydrological importance, as rivers with headwaters in the park feed the basins of the Orinoco and Amazon rivers. Picachos is a Category II Park as designated by the International Union for Natural Conservation with the stated goal of conserving large-scale ecological processes, species, and ecosystems for current and future generations (Dudley (Editor), 2008). However, more than 200 families – many fleeing armed conflict – have settled in the park over the last 25 years (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016) and introduced novel land uses such as cattle grazing and subsistence agricultural production, including illicit coca cultivation (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015). Moreover, during the early 2000s, the surroundings of Picachos were legally sited by FARC–EP (Fuerzas Armadas Revolucionarias de Colombia–Ejercito del Pueblo) as part of a negotiation process.
conducted by the former Colombian government. This agreement collapsed, and the conflict intensified from 2002 to 2010. In recognition of these cross-border processes, our study area includes Picachos as well as the 10 km region surrounding the park.

![Figure 1](image-url)

Figure 1. The study area of Picachos National Park is located in central Colombia at the northern extent of the Andes Mountain range.

2.2.2 MAIAC data

The MODIS sensor is a pillar of large-scale remote sensing studies due to its superior data quality and daily global coverage (Huete et al., 2002). MODIS surface reflectance is regularly derived from top of atmosphere measurements using pixel-based atmospheric correction and cloud screening (Vermote & Kotchenova, 2008). Nevertheless, errors in the estimation of atmospheric aerosol loadings (Grogan & Fensholt, 2013; Samanta et al., 2010; Samanta, Ganguly, Vermote, Nemani, & Myneni, 2012) and deficiencies in cloud screening (Hilker et al., 2012) introduce variability in estimated surface parameters unrelated to actual changes in land cover.
(Zelazowski, Sayer, Thomas, & Grainger, 2011), and these may lead to incorrect quantification of vegetation trends (Samanta et al., 2012).

As an alternative to conventional MODIS at-surface products, the Multi-Angle Implementation of Atmospheric Correction (MAIAC) is a novel cloud screening technique that enhances land surface reflectance based on MODIS C6 Level 1B data (Hilker et al., 2014). MAIAC’s cloud detection technique yields a noise reduction by a factor of 3-10 compared to standard MODIS surface reflectance products (MYD09, MYD09GA) and derived composites (MYD09A1, MCD43A4, and MYD13A2-Vegetation Index) without the conventional assumption of Lambertian land surface behavior (Hilker et al., 2012, 2015). The improvements offered by MAIAC increase the number of viable clear-sky observations by a factor of 2-5 (Hilker et al., 2012), which is especially valuable for monitoring vegetation dynamics in the persistently clouded tropics (Seddon, Macias-Fauria, Long, Benz, & Willis, 2016; Wilson & Jetz, 2016). We used 8-day composite NDVI observations at 1km resolution from January 2001 to June 2014 as input for the BFAST analysis.

2.3 MAIAC-based trend detection with BFAST
BFAST iteratively decomposes time series data into trend, seasonal and noise components and supports detection of structural changes (i.e., breakpoints) in both trend (i.e., browning or greening) and seasonal components (de Jong, Verbesselt, Schaepman, & de Bruin, 2012; Verbesselt, Hyndman, Zeileis, & Culvenor, 2010). The parametrization of BFAST is based on a single parameter, $h$, the inverse of which identifies the largest number of breakpoints that could be detected in a given time series (Watts & Laffan, 2014). We set $h=0.15$ to reduce the influence of persistent cloud cover that introduces missing data into the MAIAC-NDVI time series. With our MAIAC-NDVI time series being 13.5 years in duration, $h=0.15$ also means we would have a minimum segment length of at least 2.03 years (93 observations) between breakpoints.
We evaluated the ecological significance of each segment’s slope ($\beta$) output by BFAST (Figure 2). The trend within each segment is assumed to be linear and is estimated using robust linear regression (de Jong et al., 2012). Only segments with significant $\beta$ values (i.e. non-zero trends) were considered to be potential greening (positive slope) or browning (negative slope) segments, and were tested against the null hypothesis of $\beta=0$ at $\alpha=0.05$. At each pixel, the total respective durations of greening and browning periods were calculated as the sum of individual segments with significant slopes. All analysis based on MAIAC were performed in R using the greenbrown library (Forkel et al., 2013).
2.2.3 Landsat imagery

We used all available Level-1 Terrain-Corrected (L1T) surface reflectance-corrected Landsat images from path/row 8/58, which covers Picachos in its entirety. In total, this amounted to 374 Landsat 5, 7 and 8 images from 1996-2015; images with over 50% cloud cover were removed, leaving 149 Landsat images for the analysis (Table 6, Figure 17 Appendix). Cloud masks produced using Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (J.G. Masek et al., 2013), Function of Mask (FMask) (Zhu & Woodcock, 2012), and Landsat 8 Surface Reflectance (L8SR) algorithms (Holden & Woodcock, 2015; Zhu et al., 2016) were applied, and NDVI was calculated for all images. Although other vegetation indices have shown less sensitivity to saturation issues in tropical forests, recent studies have demonstrated that NDVI is less affected by sun-sensor geometry variation (Hilker et al., 2015; Morton et al., 2014) as well as variation in the frequency distribution of observations (Schultz et al., 2016), both of which are relevant for time series analysis.

The cross calibration between Thematic Mapper (TM), and Enhanced Thematic Mapper Plus (ETM+) sensors supports a continuity of cross-sensor NDVI trends during the study period (Chander, Markham, & Helder, 2009; Vogelmann, Gallant, Shi, & Zhu, 2015; Zhu et al., 2016). However, TM/ETM+ and Landsat 8 Operational Land Imager (OLI) do not offer such consistency. Zhu et al. (2016) found that OLI-derived NDVI values were positively biased with respect to ETM+ by about 0.04. Consequently, we transformed NDVI surface reflectance-corrected data using the methods described by (Roy et al., 2015) to minimize OLI’s positive bias and ensure NDVI consistency across Landsat sensors.

To simplify image processing, a forest/non-forest mask was applied to the Landsat archive by masking out all pixels with less than 50% tree cover from the “Year 2000 forest cover” data product from (Hansen et al., 2013). We also removed pixels labeled as “rainfed cropland” according to the Land Cover V 2.5 product based on 1 km SPOT-VEGETATION and 300 m Medium Resolution Imaging Spectrometer (MERIS) data with the epoch centered on 2000 (1998–2002) (Arino, Perez, Kalogirou, Defourny, & Achard, 2010). Following recommendations by (Hamunyela et al., 2016), we also masked pixels with less than 30 observations (20%
of the full time series) and less than 7 observations in the historical period (1996-2000) (see section 2.2.4 and Figure 3).

Figure 3. Example of pixel-level BFAST Monitor output using a 2007-2010 subset of the Landsat time series. Black solid lines denote the beginning of the monitoring period, black dots are observed NDVI observations, blue lines denote expected values from the fitted seasonal model, and grey dashed lines represent the residuals in each monitoring period. A breakpoint (dashed red line) is identified in early 2007 with a magnitude (M) of -0.35 but is not detected in 2008 or 2009. The historical periods for 2008-2009 and 2009-2010 are shorter than the 2007-2008 period because BFAST Monitor only includes preceding observations until the stable seasonal model breaks down.

2.2.4 Landsat-based disturbance detection with BFAST Monitor

Using the Landsat time series, we applied BFAST Monitor to map forest disturbance in Picachos given its proven robustness at detecting disturbance in other tropical forests (Schultz et al., 2015). BFAST Monitor differs from the conventional BFAST algorithm that we used with the MAIAC time series in three important ways. First, we fit only a first-order harmonic curve to the over Landsat time series using BFAST
Monitor, whereas we used a third-order harmonic for the MAIAC time series. This decision lies in the irregular distribution of Landsat observations that can result in model over-fitting (DeVries et al., 2015). Second, we omitted the trend component when modeling the Landsat time series but used a seasonal-and-trend model with MAIAC. The inclusion of a trend component in modeling the Landsat time series would have generated erroneously modeled NDVI values in the monitoring period. Third, BFAST extracted major breaks by considering the MAIAC time series as a whole, while BFAST Monitor evaluated potential breaks across the Landsat time series in a succession of one year long periods.

We used 32 Landsat images collected between 1996-2000 to develop a historical baseline of forest cover trends. Because using all images may affect the BFAST Monitor model fit (Verbesselt et al., 2012), we measured the reversed-ordered-cumulative (ROC) sum of residuals to generate a cumulative prediction error that helps identify whether a seasonal model no longer offers an accurate fit (Pesaran & Timmermann, 2002) (Figure 3). With the historical model in place, we applied BFAST Monitor sequentially from 2001 to 2015, following the approach by DeVries et al. (2015).

The detection of breakpoints using BFAST Monitor is associated with near-zero or positive magnitudes which could not be related to forest disturbance (DeVries et al., 2015). The breakpoint magnitude is calculated as the median of observation-model residuals within each monitoring period (Verbesselt et al., 2012). In identifying the magnitude threshold to distinguish disturbance from non-disturbance and minimize spurious breakpoints, we selected a stratified random sample of 420 breakpoints based on three magnitude quantiles from the study area where very high resolution images from Google Earth were available for 2004 and 2009. We classified sample points as being disturbed or non-disturbed based on visual interpretation of true color imagery as well as NDVI time series. We then calculated a magnitude threshold of -0.041 using Binomial Logistic Regression where the likelihood of actual disturbance approached 50% (Figure 18 Appendix) (Devries & Verbesselt, 2015). We considered any potential breakpoint with a magnitude less than -0.041 to be disturbed (Milton, Fox, & Schaepman, 2006).
2.2.5 Validation and agreement assessment

BFAST Monitor validation approaches in tropical regions typically involve a single image, which makes it difficult to capture annual dynamics or disturbed areas (DeVries et al., 2015; Dutrieux et al., 2015). To overcome this limitation, we used sub-annual Landsat L1T-derived NDVI imagery, Google Earth-hosted annual Landsat true color images and very high resolution images (2003-2004, 2007 and 2009) and RapidEye images (2014-2015). We visually inspected disturbance sites and identified obviously erroneous disturbances, such as pixels influenced by river dynamics.

With erroneous disturbance sites removed, we sampled 684 validation sites following Cochran’s recommendations (Cochran, 1977). After removing 24 sample sites due to lack of images, we distributed the remaining 660 sites into disturbed and non-disturbed sample sets following Olofsson et al. (Pontus Olofsson et al., 2014); this yielded 269 disturbed and 391 initial non-disturbed sites. Given that the amount of disturbed forest in Picachos is expected to be relatively small compared to stable forest or otherwise non-disturbed sites, we sought to decrease the potential overrepresentation of non-disturbed sites in the validation sample by stratifying non-disturbed validation sample with respect to a 2km buffer from unpaved roads (2014-2015) derived from high-resolution images (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016) with the assumption that disturbed sites are likely closer to roads (Barber, Cochrane, Souza, & Laurance, 2014; Bax, Francesconi, & Quintero, 2016). Our final validation sample includes 151 non-disturbed sites within the road buffer, and 240 non-disturbed sites outside the buffer.

For each disturbed site, we evaluated the temporal agreement of BFAST Monitor-derived disturbances against our reference datasets (Google Earth, RapidEye and NDVI sub-annual scenes) using 20 sampled pixels per year into TimeSync (Cohen et al., 2010) as adapted in R by (DeVries et al., 2015). A pixel-count confusion matrix was derived from every strata, and we removed the bias effect of unequal omission and commission errors by constructing an area-adjusted error matrix (Pontus Olofsson et al., 2014; Pontus Olofsson, Foody, Stehman, & Woodcock, 2013). We calculated the stratified estimator for a given class as:
\[ \hat{A}_j = \hat{A}_{total} \times \hat{p}_j \]  

(1)

Where \( \hat{A}_j \) is the adjusted area, \( \hat{A}_{total} \) is the total mapped area for a given class, and \( \hat{p}_j \) is the column sum of the cell area proportion in the error matrix. We further validated MAIAC trends using Landsat-based disturbance events by comparing the timeline of MAIAC-based breakpoints with annual Landsat-based disturbances.

2.3 Results

2.3.1 MAIAC-derived browning and greening patterns

The most prominent browning region is in the southeastern corner of Picachos with small and scattered browning regions beyond the park's northeastern perimeter near small settlements (Figure 4a). The watershed most affected by persistent browning is the Platanillo River watershed while the Templado River watershed is only affected at its southern extent where it merges with adjacent watersheds. The Yulo River watershed presented both browning and greening over our study period with more greening evident within Picachos. Greening clusters (Figure 4b) were found on a steeply sloped region in the park's central-south region and surrounded by a river, the Rio Pato, and also in the central-eastern foothills along the large rivers, the Rio Guayabero and Rio Leiva. Within the park, 12,500 ha showed 13 years of consecutive browning, whereas the total area of persistent greening amounted to 27,700 ha.

Each pixel may exhibit greening or browning at different periods over the time series (Figure 4). To calculate the percentage of Picachos that experienced browning or greening, we divide the total amount of browning and greening pixels by Picachos' extent as well as the 10km-wide region surrounding the park. In total, 14% and 20% of the park displayed browning and greening, respectively, while 32% of Picachos was masked due to absence of MAIAC observations due to cloud cover. Within the 10km wide region surrounding Picachos, 5% and 2.5% of the surrounding forest exhibited browning and greening, respectively. Therefore, we recorded a total
positive (greening) net trend of 6% within Picachos, but a negative (browning) trend of 2.5% surrounding the park.

The lengths of browning periods were generally shorter than greening periods: browning periods frequently lasted 3 or 6 years, while greening tended to last 7 or more than 10.5 years (Figure 4c). Both greening and browning trends occurred at a low elevation between 300-600m and during 2004-2007.

Figure 4. Duration of MAIAC-based a) browning and b) greening trends based on the number of years with significant trends ($\alpha=0.05$), and (c) distribution of browning and greening periods across Picachos and its surrounding area.

Fifty percent of pixels in the park had MAIAC trends with zero breakpoints over the 2001-2014 study period, 12% of the park saw a single breakpoint, while 4.5% of the park saw two or more breakpoints during the study period, most of which occurred in the southeastern corner of the park.
2.3.2 Landsat-based disturbance distribution

Using BFAST Monitor with Landsat data, we measured a total of 6353 ha of disturbed area within Picachos from 2001–2015 and 6289 ha disturbed area between 1998–2001, yielding a total disturbed area of 12,642 ha (±1440 margin of error, 95% confidence interval) or 4.3% of Picachos’ total area. This disturbed area estimate is 20% less than the previous estimate of 15,276 ha based on the CLC methodology (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016). This difference can be explained by the poor validation of the CLC dataset in this region, differences in cloud-masking approaches, and, further, that inter-annual change maps based on CLC datasets suggest forest cover change in many regions despite the presence of visually interpreted stable forest.

Figure 5. Cumulative BFAST Monitor-derived disturbance based on Landsat 2001-2015 time series, aggregated to 1km pixel size for display.

As with the spatial distribution of MAIAC browning trends, Landsat disturbances were concentrated in the southeast with scattered disturbance recorded in the northeast just outside Picachos (Figure 5). The mean size of Landsat-derived
disturbed patches is 1.9 ha (range 1.1–2.9 ha) across all years, while the maximum annual mean patch size is 19.8 ha (range 5.1–49.3 ha). The size of disturbed patches within Picachos is considerably larger than the 0.6 ha patch size in other tropical forest regions (DeVries et al., 2015), but is smaller than the 8–10 ha patch size in other mountainous regions (Fernández et al., 2015). During 2001–2015, the total disturbance area within the 10 km-wide region surrounding Picachos was 9218 ha, which is 1.6 times greater than the total amount of disturbance within the park. The years 2004 and 2007 witnessed the most disturbance in the region surrounding Picachos, whereas 2007 and 2014 exhibited the most change within the park (Figure 6).

![Figure 6. Temporal distribution of disturbance area within Picachos and the surrounding region. The year 2003 was not evaluated because of a lack of Landsat images.](image)

2.3.3 Validation of Landsat-derived disturbances

Using 660 validation points, we recorded a high total accuracy (TA) of 0.99±0.007 for disturbed class sites, with user’s (UA) and producer’s accuracies (PA) of
0.95±0.024 and 0.83±0.18, respectively. The non-disturbed classes presented a combined UA of 0.98±0.026 and PA of 0.99±0.002 (Table 1). Accuracies were similar to previous studies in South America (Devries & Verbesselt, 2015), but higher in comparison with other tropical regions (Schultz et al., 2016). We also assessed the agreement between Landsat-based BFAST Monitor disturbance dates and TimeSync-interpreted disturbance dates based on sub-annual Landsat true color, sub-annual Landsat L1T-derived NDVI scenes and supporting very high resolution imagery. Most disturbance dates showed an agreement of 75% within ± 6 months, while the bulk of other disturbances (22%) showed a residual of 6-12 months; and 3% of remaining errors showed disagreement greater than 12 months (Figure 7).

Table 1. Area-weighted error matrix of Landsat-based BFAST Monitor results using stratified sample sites across classes: 1) Disturbance, 2) Non-disturbance outside 2km unpaved roads buffer, and 3) Non-disturbance inside 2km unpaved roads buffer.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Proportion of area mapped (Wi)</th>
<th>UA</th>
<th>PA</th>
<th>TA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.027 0.0001 0.0011</td>
<td>0.028</td>
<td>0.95±0.024</td>
<td>0.83±0.18</td>
</tr>
<tr>
<td>2</td>
<td>0.002 0.172 0</td>
<td>0.174</td>
<td>0.98±0.018</td>
<td>0.99±0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.003 0 0.793</td>
<td>0.797</td>
<td>0.99±0.008</td>
<td>0.99±0.001</td>
</tr>
<tr>
<td>Total</td>
<td>0.033 0.172 0.794</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.3.4 Validation of MAIAC breakpoints

The most frequent year of MAIAC breakpoints was 2007, followed by 2004 and 2010, which coincide with the most frequent years of Landsat disturbance between 2002–2013 (Figure 8). Although MAIAC time series data spanned from January 2001 to June 2014, no significant breaks were detected during 2001–2002 and 2013–2014. This is due to the h-defined minimum segment size of 2.03 years that restricted our ability to detect breakpoints at the beginning and end of the time series but supported accurate model fitting within each segment (De Jong, Verbesselt, Zeileis, & Schaepman, 2013; Verbesselt et al., 2012; Watts & Laffan, 2014). The pattern of Landsat disturbances is more temporally-resolved because of the year-by-year sequential monitoring used by BFAST Monitor. Nonetheless, the agreement in timing and location of major changes observed in MAIAC and Landsat time series confirms the potential of coarse spatial resolution time series to accurately capture the broad-
scale multi-year trajectory of forest disturbance.

Figure 8. Comparison between years of breakpoints detected using MAIAC and Landsat-based disturbances analyses in whole study area.

2.4 Discussion

In this study, we demonstrated the ability of BFAST-based methods to accurately detect short- and long-term disturbances over a mountainous forest region. This is also the first study to combine all available Landsat 5, 7, and 8 images using BFAST Monitor and account for radiometric and spectral differences between TM/ETM+ and OLI sensors; while (Devries et al., 2016), used the same sensors within BFAST but did not consider the potential bias in cross-sensor NDVI measurements (Roy et al., 2015; Vogelmann et al., 2015). Accounting for this bias is critical since the bias magnitude may approach or exceed the threshold selected to identify disturbance (Devries & Verbesselt, 2015; DeVries et al., 2015; Hamunyela et al., 2016).
2.4.1 Possible drivers of vegetation trends and disturbances

Picachos and its surrounding region have been affected by internal conflict and population displacement since as early as 1991 (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016). During 1998-2002, FARC-EP (Fuerzas Armadas de Colombia- Ejercito del Pueblo) was legally sited in San Vicente del Caguán-neighboring Picachos to the south, as part of a negotiation process conducted by the former Colombian government. However, this agreement collapsed and the increased violence associated with the conflict between FARC and the Colombian government has been widely recognized as generating both internal displacement and economic migration into Picachos, particularly in the accessible southeastern region (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016; Vásquez, 2013). The increased in-migration accelerated the expansion of unpaved roads and infrastructure (small settlements) as well as selective logging and clearing for crops and cattle pasture, all of which has contributed to deforestation in the park (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016). Such changes in land cover and land use have been documented in a peasant reserve close to Picachos’ southern boundary, Losada-Guayabero, which includes more than 2000 families with agricultural and pastoral livelihoods (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016). The changes in dominant land use coincides with the concentration of browning trends and small-scale disturbances in the southeastern foothills of Picachos as detected by both MAIAC and Landsat and corroborated in previous studies (D. Armenteras et al., 2003; Dolors Armenteras et al., 2011; Etter, McAlpine, et al., 2006; Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016) (Figure 9). Conversely, the concentration of MAIAC-based greening trends and absence of Landsat-based disturbance near large rivers in eastern Picachos, such as Rio Guayabero, suggests that these rivers act as barriers to forest resource extraction.

The abundance of disturbance events between 2004 and 2010 is likely explained a consequence of increased beef consumption in Colombia in the mid-2000s (Federacion Nacional de Ganaderos de Colombia (FEDEGÁN), n.d.) and the growth of the livestock industry from 20 to 23 million head during 2001-2010 before stabilizing over 2011-2014 (Jimenez, Miranda, & Gantiva, 2008). Furthermore, Picachos is sited in the provinces of Meta and Caquetá, provinces that saw a high growth of livestock production to supply the domestic market (Santos, 2015), which has encroached upon Picachos and its surroundings.

The increase in disturbance surrounding Picachos had its roots in agrarian policies that have contributed to the intensification of agricultural land uses over lowlands and foothills within the region (Vásquez, 2013). These policies promoted an initial process of occupation and accumulation of land followed by the establishment of production models based on agricultural and livestock-based land use systems.
Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015). The changes brought by these formal land use policies were subsequently exacerbated by other regional processes such as: 1) expansion of illegal crops, 2) forced displacement due to armed conflict, and 3) lack of records of land titles and hence land tenure, which have often resulted in an informal and unrecognized appropriation of land (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016). While a broader discussion of socio-economic drivers is warranted, such an examination is made difficult by lack of field and demographic data, as well as accessibility to the study site.

2.4.2 Effectiveness of BFAST disturbance detection methods

Both BFAST implementations supported detection of vegetation change with high spatial and temporal accuracies. Though the MAIAC time series has a coarse spatial resolution, its high temporal resolution, improved cloud detection technique, and exceptional mitigation of sun-sensor geometry supported a refined detection of breakpoints that were broadly confirmed using a year-to-year Landsat-based disturbance monitoring approach. The most abrupt changes in the Landsat time series were recorded in 2004, 2007, 2010 and 2014. Our detection of such abundant changes despite the low average of 8 Landsat images per year from 2001-2015 suggests that having a relatively low number of images during the monitoring period had little effect on our ability to detect disturbances (See Table 6 Appendix).

The Landsat-based disturbance class presented a high spatial and temporal agreement with the majority of errors (22%) resulting from a 6-12 month premature assignation of disturbance. Omission errors (17%) could have been improved by increasing the density of the time series (Schultz et al., 2015), whereas commission errors (5%) could have benefitted by stratifying validation site samples according to likelihood of vegetation change, such as locations along the park border, along unpaved roads, or near rivers.

The reference datasets employed showed agreement in disturbance timing at over 75% of sampled sites when compared to our TimeSync-based interpretation. NDVI also performs well despite its well-known saturation issues over areas with
high biomass values. One reason NDVI was useful is likely because the rapid alteration of Picachos was more readily detected than gradual or low-intensity change disturbances over time. Had this not been the case, the NDVI performance may have suffered, prompting us to pursue another vegetation index (Schultz et al., 2015). Other approaches such as Spatial Temporal Adaptive Algorithm for mapping Reflectance Change (STAARCH) (Hilker, Wulder, Coops, Seitz, et al., 2009) or the Noise Insensitive Trajectory Algorithm (NITA) (Alonzo et al., 2016) would also have been useful in estimating trends over these mountainous and persistently cloudy places.

2.4.3 Other Limitations

This study provides new insights into the contribution of small-scale disturbances and long-term trends over a 15-year period in one of the most inaccessible national parks in the Andes. While we have recorded short- and long-term changes in Picachos using multi-sensor time series data with high spatial and temporal agreement, there are three main areas for improvement in future studies. Disturbance detection was not possible over 32% of Picachos, primarily at the highest elevations where cloud cover was near-constant. While we expect there to be minimal human-induced disturbance at elevations over 2000 meters above sea level (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016), the lack of cloud-free observations means we cannot examine potential climate change-related vegetation degradation. As with other studies (DeVries et al., 2015; Schultz et al., 2015), we also do not specifically identify vegetation recovery and assumed that a BFAST Monitor-detected breakpoint with a magnitude less than our threshold (-0.041) symbolized a permanent change. Although this assumption is reasonable for land cover conversion caused by land use changes such as livestock management and road expansion, it may not be appropriate for detecting natural disturbances. Furthermore, we do not make any differentiation between deforestation and degradation. Though NDVI is sensitive to significant change in vegetation condition, NDVI alone may not support detection of forest thinning or degradation (Devries et al., 2016).
Greening and browning indicators from MAIAC are only represented in terms of individual significant slope ($\beta$) segments. Although the magnitude of change may have value for classifying the severity of disturbance (i.e. gradual or abrupt), the majority of magnitude values were small: browning magnitudes ranged from -0.318 to -0.00072 (mean=-0.055), whereas greening ranged from 0.0005 to 0.239 (mean=0.05). These values must be interpreted with caution due to pixel size and NDVI saturation issues over dense forest.

Although we reached high disturbance detection accuracies, omission errors suggest that data availability is still a problem. As with many other disturbance analyses in tropical regions, our lack of Landsat coverage during the 1990s affects our ability to continuously monitor long-term changes (Broich et al., 2011; DeVries et al., 2015). In our case, zero Landsat images are available between 1992-1995. Given this, it would be valuable to include observational data such as synthetic aperture radar (Reiche, 2015), Sentinel 1-2 (Petrou, Manakos, & Stathaki, 2015), or very high resolution commercial imagery from the Planet Labs, for example (Houborg & McCabe, 2016). A related limitation stems from the masking procedure combining Hansen et al.(Hansen et al., 2013) and European Space Agency Corine land cover products, neither of which are well validated for this study area.

Finally, we do not perform a quantitative attribution of change events to specific agents of disturbance. Of special interest, the impact of armed conflict over land use changes and internally displaced people (IDPs) along with illegal land appropriation has not been sufficiently explored, nor have the effects of forced displacement on the potential for forest recovery. Investigation of such topics would make for a promising future study.

2.3 Conclusions

In this study, we recorded short-term disturbances and long-term trends of forest disturbance in Picachos National Park within the Colombian Andes from 2001-2015. We implemented and validated two approaches based on BFAST, neither of which had been tested over persistently cloudy regions. Our BFAST implementations using MAIAC and Landsat identified greening and browning periods based on significant
trend segments as well as smaller disturbances with relatively high spatial and temporal accuracies. We identified an abundance of greening within Picachos, albeit with a hotspot of vegetation decline in the park’s southeastern foothills region that likely results from population incursion and land use conversion. This study was limited by a lack of MAIAC and Landsat observations at higher elevations, which affects both our long-term evaluation of forest trends as well as evaluation of short-term ecological processes. Although MAIAC offers a coarse spatial resolution, our results show that MAIAC-based time series can effectively detect hotspots of vegetation decline where Landsat data are not available.

2.4 References


Chapter 3. Monitoring forest disturbance and attributing drivers of change in the Colombian Andes

3.1 Introduction

Colombia is the second most biologically diverse country on Earth, containing 14% of the world’s biodiversity and home to approximately 10% of the world's species (Butler, 2006). Furthermore, it is the third most populous country in Latin America with over 48 million people. Its economy is based on agriculture, industry, and services (manufacturing, construction, mining, telecommunications) (Sanchez-Cuervo & Aide, 2013). The quick expansion of human population and the growing demand for agricultural land have increased the pressure on Colombia’s forest ecosystems (Etter, Mcalpine, Phinn, Pullar, & Possingham, 2006). Though Colombia has included 12% of the national land mass within a Protected Areas, (Sanchez-Cuervo & Aide, 2013) they remain threatened by drivers associated with population growth, rural-urban forced migration and economic development (Dolors Armenteras et al., 2011; Etter, McAlpine, et al., 2006; Sánchez-Cuervo & Aide, 2013). As a result, a disturbance pattern has been detected over the Andes–Amazonia transition belt (Achard et al., 2014; Dolors Armenteras et al., 2011) and specifically in the Colombian Andes foothills (Etter et al., 2008; Stibig, Achard, Carboni, Raši, & Miettinen, 2014) that corresponds to locations of increasing anthropogenic pressure since the early 1950s. The foothills of the Colombian Andes have also been affected by a long history of internal social conflict, forced migration, and illicit cultivation of coca (*Erythroxylum coca*), with attendant effects on socio-ecological relationships (Dolors Armenteras et al., 2011; Sanchez-Cuervo & Aide, 2013; Sánchez-Cuervo, Aide, Clark, & Etter, 2012; Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016).

Previous studies within Colombia have identified causes of forest cover change at national scales through comparisons of two or three Landsat imagery separated from each other by five or more years (Etter, Mcalpine, et al., 2006; Rodríguez et al., 2013). Other approaches have employed more regularly collected images derived from MODIS, but are based on either measuring the effectiveness of Protected Areas at limiting deforestation (Cortes Alonso & Sergio, 2012; Salazar
Villegas, 2013) or detecting main drivers of forest cover change at the municipality-level (Sanchez-Cuervo & Aide, 2013). As a consequence, there are no available studies that evaluate small-scale drivers of change at high temporal resolution in the Colombian Andes foothills.

A diversity of metrics derived from remote sensing time series data can be used to indicate, not only the presence of forest cover change, but also the drivers of change. LandTrendr (Kennedy et al., 2015), the vegetation change tracker (VCT, (Huang et al., 2010), and trend detection approaches based on “best-available-pixel” (Frazier, Coops, Wulder, & Kennedy, 2014; Hermosilla et al., 2015; Pickell et al., 2014) have all been commonly used for disturbance detection. However, these approaches produce annual disturbance maps, which may miss subtle, short-lived, or otherwise sub-annual changes (Forkel et al., 2013). These analyses are also reliable in regions with less frequent cloud cover, a deeper Landsat archive (>30 years), or with more regular data availability, e.g. United States and Canada (Hermosilla et al., 2015; Kennedy et al., 2015; Pflugmacher et al., 2012; Pickell et al., 2014).

Using 149 Landsat scenes during the period 1996-2015, we mapped disturbances using Breaks For Additive Season and Trend (BFAST) algorithm and assigned the most probable drivers of change in Picachos National Park, focusing on recognized drivers of change i.e. conversion to pasture, conversion to subsistence agriculture and non-stand replacing disturbance. We initially built a sample of common drivers using high-resolution imagery and Corine Land Cover (CLC) maps. We characterized each driver using spectral metrics before and after disturbance detection, which describe persistence or recovery of vegetation, as well as conventional trend (e.g. magnitude and slope), pattern, and topographic metrics.

Trend, spectral, pattern, and topographic metrics as well as training data on sites and drivers of change were input to a Random Forest classifier to attribute the most probable driver of change. Revealing the patterns of forest cover change as well as associated drivers will inform the adoption of sustainable strategies in areas like Picachos with high biological diversity but affected by unregulated and often undocumented human encroachment.
3.2 Study area

The study area is located in the foothills of the Colombian National Park, Picachos, and covers the basins of the Templado, Platanillo, and Yulo rivers. Picachos has an elevation gradient between 300-2000masl, and an annual accumulated rainfall varying from 1050-3250mm (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016). Picachos is a Category II Park as designated by the International Union for Conservation of Nature (IUCN) with the stated goal of conserving large-scale ecological processes, species, and ecosystems for current and future generations (Dudley (Editor), 2008). Although people are not legally allowed to reside within the park, more than 200 families and 119 small houses and plots with sizes among 50 to 250 ha have been recorded within the park using Corine Land Cover (CLC) datasets, and these communities have been responsible for expanding grazing lands and subsistence agricultural production (Sánchez-Cuervo & Aide, 2013; Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016; Vásquez, 2013).

Figure 10. Study area located over Picachos National Park foothills
3.3 Methodology

As an overview of our approach for relating disturbance events to drivers, we first conceptually divided the study area into two broad classes, stable forest (no disturbance) and disturbance based on BFAST Monitor disturbance outputs (Figure 11). Pixel-level disturbance were aggregated to patches based on Rook’s case adjacency, and only patches with an area greater than 1 ha (~ 9 pixels) were considered in the analysis. Disturbance patches were then conceptually categorized as long- or short-term disturbance. Long-term disturbance may include conversion to pasture or agriculture whereas short-term disturbance are non-stand replacing that may result from insect infestation, disease (Hilker, Coops, et al., 2009), stressed vegetation (Hermosilla et al., 2015), low-severity fires (Kennedy et al., 2015), as well as human-induced activities such as low-intensity logging (Souza, Roberts, & Cochrane, 2005). Using high-resolution imagery and CLC maps we developed a training database for the most common drivers in Picachos. We then used spectral, trend, pattern, and topographic metrics to associate specific drivers to mapped disturbance patches using a Random Forest (RF) classifier (Breiman, 2001), and summarized changing driver dynamics over the course of the study period.

Figure 11. Conceptual flowchart for disturbance detection and driver attribution.

3.3.1 Image pre-processing and NDMI calculation
Previous studies have shown that typical patch sizes in Picachos range from 1-50ha (mean=2ha) (Murillo-Sandoval et al., 2017). To detect such small scale changes, we used all available Level-1 Terrain-Corrected (L1T) surface reflectance-corrected Landsat images from path/row, 8/58. A total of 149 images were used for this analysis from 1996 to 2015 after removing 225 images with more than 50% cloud cover. Cloud masks were produced using Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS, Masek et al., 2013), Function of Mask (FMask, Zhu & Woodcock, 2012) and Landsat 8 Surface Reflectance (L8SR) algorithms (Holden & Woodcock, 2015; Zhu et al., 2016). Since BFAST requires identifying a period of stable forest cover for proper calibration, we developed a baseline forest cover mask using a supervised classification of cloud-free images from 1996-1998.

NDMI, also known as Normalized Difference Moisture Index (Gao, 1996) combines short-wave infrared (SWIR) and near infrared (NIR) to identify canopy moisture content (Hayes & Cohen, 2007) and has been successfully used to monitor disturbance and regrowth in tropical places (Devries & Verbesselt, 2015). NDMI has shown better sensitivity for disturbance detection than greenness indices such as EVI and NDVI in tropical forests (Schultz et al., 2016), and is effective even with intra-annual variability in image availability, a limitation common across the Andes region (Schultz et al., 2016).

\[
NDMI = \frac{NIR-SWIR}{NIR+SWIR}
\]  

The simplicity and performance of NDMI over tropical places has demonstrated similar mapping accuracies for disturbances in comparison with other metrics that exploit the full dimensionality of all Landsat bands such as Tasseled Cap Wetness (TCW) (Jin & Sader, 2005) and Normalized Difference Fraction Index (NDFI) which involves spectral mixture analysis using fraction of green vegetation, shades and soils components (Schultz et al., 2016; Souza et al., 2005).

3.3.2 Disturbance detection using BFAST Monitor time-series analysis

Given its proven robustness for monitoring changes across various tropical forests, the BFAST Monitor algorithm was select to successively map disturbances (DeVries
et al., 2015; Dutrieux et al., 2015; Hamunyela et al., 2016). BFAST incorporates the iterative decomposition of time series data into trend, seasonal, and noise components, with methods for detecting changes (Verbesselt, Hyndman, Newnham, et al., 2010; Verbesselt et al., 2012). We applied BFAST Monitor in 1-year sequential monitoring periods beginning in 1999 using a first-order harmonic curve over NDMI time series. We also omitted the trend component to avoid unrealistic NDMI values in the monitoring period which caused the detection of false breakpoints (DeVries et al., 2015).

Breakpoints detected by the BFAST Monitor are robust against noise and other variability in the time series (Verbesselt et al., 2012). However, a magnitude threshold was necessary to avoid the detection of spurious breakpoints due to outliers in the monitoring period (DeVries et al., 2015). Using a visual comparison against a WorldView-1, 0.5m resolution image from December 12, 2009, we selected an NDMI change magnitude threshold of -0.01; all pixels with a change magnitude less than this threshold were classified as disturbed.

3.3.3 Disturbance validation using TimeSync

A stratified sampling procedure was used to validate the presence of changes (Pontus Olofsson et al., 2014). Using equation (5.25) (Cochran, 1977) and assuming a target standard error for overall accuracy of 1%, a total of 614 pixels were selected in two strata change and no change; 22 pixels were removed because of unreliable validation due to excessive cloud cover in the annual Landsat imagery and high resolution reference datasets. We sampled 211 pixels for change strata and 381 pixels for no change strata. A visual interpretation of changes was undertaken with TimeSync (Cohen et al., 2010; DeVries et al., 2015) using annual mean Landsat NDVI composites and Google Earth and RapidEye high-resolution images (P Olofsson, Holden, Bullock, & Woodcock, 2016). A confusion matrix was created by cross-tabulating the BFAST-derived disturbance map and reference imagery datasets, and estimated areas of disturbance with confidence intervals were constructed with stratified estimation.
3.3.4 Characterizing disturbance drivers

The disturbance events identified by BFAST Monitor involve two components: the change date and the change magnitude at patch-level. In cases that a driver and its magnitude of change are well characterized in a given study area, the disturbance event can be attributed to a specific driver (Reents, 2016). However, since a similar change magnitude may be common across drivers, e.g., the conversion of stable forest into pasture may show the same magnitude of change as a conversion to agriculture, other topographic, geometric pattern, and spectral trend metrics are useful for improving the identification of specific potential drivers (Kennedy et al., 2015; Pickell et al., 2014). For example, *conversion to agriculture* typically presents an abrupt change and post-disturbance spectral index cycling that indicates vegetation growth. *Non-stand replacing disturbance*, on the other hand, shows subtle negative slope over time, and *conversion to pasture* presents an abrupt break followed by a sustained period of relatively low vegetative condition (Figure 12).

![Figure 12](image)

Figure 12. Example BFAST Monitor-derived disturbance trajectories associated with drivers of change. Vertical red lines indicate a breakpoint, blue lines indicate a seasonal+trend model fitted and M denotes the magnitude of change.
We calculated the average and standard deviation for near infrared (NIR), shortwave infrared 1 (SWIR1), shortwave infrared 2 (SWIR2) and NDMI before and after the date of disturbance (Healey, Cohen, Zhiqiang, & Krankina, 2005; Hermosilla et al., 2015). As a measure of change severity across time, we measured the mean pixel-level NDVI at each disturbance pixel for each year and calculated the trend using ordinary least squares (OLS) regression (Forkel et al., 2013). Pixel-level trends were then averaged across each patch (Table 1). We also derived pattern metrics including patch area, perimeter (Bogaert, Rousseau, Van Hecke, & Impens, 2000), shape index (McGarigal, Tagil, & Cushman, 2009), and fractal dimension (Krummel, J. R., Gardner, R. H., Sugihara, G., O’Neill, 1987); and topographic metrics including elevation, slope, and topography position index defined as the difference between the value of a cell and the mean value of its 8 surrounding cells (Kennedy et al., 2015). Finally, we measured patch-level kurtosis and skewness for both the NDMI magnitude of change and NDVI trend (Pickell et al., 2014).

Table 2. Summary of patch-level metrics used to characterize drivers of disturbance. SRTM = Shuttle Radar Topography Mission Digital Elevation Model (90m).

<table>
<thead>
<tr>
<th>Topographic indicators</th>
<th>SRTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>SRTM</td>
</tr>
<tr>
<td>Slope</td>
<td>SRTM</td>
</tr>
<tr>
<td>Topographic position index</td>
<td>SRTM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pattern metrics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td></td>
</tr>
<tr>
<td>Perimeter</td>
<td></td>
</tr>
<tr>
<td>Shape index</td>
<td></td>
</tr>
<tr>
<td>Fractal dimension</td>
<td></td>
</tr>
</tbody>
</table>

| Trend time-series metrics              |      |
| Aggregated annual trend (mean, median) | NDVI |
| Average and median change magnitude    | NDMI |
| Standard deviation change magnitude    | NDMI |
| Kurtosis and skewness change magnitude | NDMI |

| Spectral summary metrics               |     |
| Average spectral value pre-change      | NIR,SWIR1,SWIR2,NDMI |
| Average spectral value post-change     | NIR,SWIR1,SWIR2,NDMI |
| Standard deviation spectral value pre-change | NIR,SWIR1,SWIR2,NDMI |
| Standard deviation spectral value post-change | NIR,SWIR1,SWIR2,NDMI |
3.3.5 Attribution of changes

Cataloging past disturbances is challenging mainly due to missing historical “ground truth” reference datasets (Cohen et al., 2010; Kennedy et al., 2015). Therefore, we used a collection of complementary maps generated by Parques Nacionales Naturales de Colombia, Instituto Geografico Agustion Codazzi and Instituto Amazonico de Investigacion Cientifica using Corine Land Cover methodology (CLC) for the years 2002, 2007, and 2012 (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2015, 2016). CLC maps provide the most reliable source for interpreting land-conversion towards pasture and agriculture but have not been validated yet for this region. We also used high resolution (0.65m) images hosted by Google Earth (Jan 17, 2003; Dec 31, 2009) and RapidEye imagery (Oct 23, 2014; Dec 15, 2014; Jan 5, 2015) to visually confirm and assign the presence of all three major drivers in our study area.

From the total population of disturbance patches (3450) we visually interpreted 476 patches and categorized each into conversion to pasture, conversion to agriculture, or non-stand replacing disturbance using CLC and support imagery. Since there was little variation in patch area (i.e., 2.4-3.9ha), a stratified selection of objects by area was not performed. The full suite of topographic, time series, pattern, and spectral metrics were measured at disturbance sites and used for training (60%) and validating (40%) a RF classifier, which has been regularly used for the classifying drivers (Devries et al., 2016; Hermosilla et al., 2015; Kennedy et al., 2015; Reents, 2016). With 36 input metrics, we limited the potential for multicollinearity by removing 9 metrics with Pearson’s R values greater than 0.70 (Hermosilla et al., 2015).

We used the majority of RF votes for driver attribution. Drivers attributed to each disturbance patch using RF were compared with the validation sample (reference) to measure the attribution accuracy. We also calculated the distance to the second most voted class as a measure of classification confidence, calculated as the ratio (d2/d1) where d2 corresponds to second most voted class and d1 is the assigned class (Mitchell, Remmel, Csillag, & Wulder, 2008). Ratio values range from nearly 0 (when d2<d1) to 1 (when d2=d1). Patches with confidence values higher than 0.6
were labeled as unclassified because of perceived unreliability of driver attribution (Hermosilla et al., 2015).

3.4 Results

3.4.1 Disturbance detection performance

The total disturbance area within the Picachos foothills was 10717±1095ha while stable forest was recorded across 53064±1094ha. Using visual interpretation results at 594 disturbance and stable forest pixels, we estimated a total accuracy (TA) of 0.94±0.02. The disturbance class had user’s (UA) and producer’s accuracies (PA) of 0.98±0.02 and 0.73±0.07, respectively; the stable forest class had UA of 0.94±0.02 and PA of 0.99. Disturbance pixels had much lower commission (2%) than omission errors (27%) whereas stable forest pixels were more affected by commission (6%) than omission (1%). Both stable forest and disturbance presented high agreement, revealing that forest disturbance and stable forest conditions can be well-detected using NDMI. We found much lower disturbance than existing CLC maps that recorded 12128ha of forest cover disturbance during a similar same study period. This difference is explained by the lack of accurate validation data such that CLC identifies forest cover disturbance in many regions despite visually interpreted forest cover stability.

Table 3. Area-weighted error matrix of NDMI-based BFAST Monitor assessment using stratified sample sites across classes: disturbance and stable forest.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Disturbance</th>
<th>Stable Forest</th>
<th>Wi</th>
<th>User’s Accuracy</th>
<th>Producer’s Accuracy</th>
<th>Total Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disturbance</td>
<td>0.134</td>
<td>0.004</td>
<td>0.14</td>
<td>0.96±0.024</td>
<td>0.79±0.08</td>
<td>0.96±0.01</td>
</tr>
<tr>
<td>Map</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable forest</td>
<td>0.03</td>
<td>0.82</td>
<td>0.86</td>
<td>0.96±0.02</td>
<td>0.99±0.004</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.186</td>
<td>0.811</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4.2 Characterization of disturbances

Variables were ranked by the RF variable importance measure (Figure 13). The mean post-change NDMI was the most important metric, followed by post-change standard deviation NDMI, NDVI trend, and mean pre-change NDMI. If these metrics were excluded from the RF model, accuracy would decrease by 33%, 21%, 15% and 13%, respectively. Other relevant but less important metrics include mean post-change NIR (11%), standard deviation of NDVI trend (8%), elevation (7.5%) and median magnitude of NDMI (4.6%). Overall, spectral metrics had the most predictive power followed by topographic and pattern metrics.

![Figure 13. Most important metrics for discrimination among drivers based on mean decrease accuracy in RF model](image)

*Conversion to pasture* is well distinguished from the other two drivers especially in pre- and post-change NDMI (Fig 14a and 14c) and post-change SWIR2 (Fig 14b), less so for elevation (Fig 14f). *Conversion to agriculture* and *non-stand replacing disturbance* showed less spectral separability, however non-stand replacing presented the smallest standard deviation NDVI trend (Fig 14d) and a range of NDMI magnitude less abrupt than the rest of drivers (Fig 14e).
3.4.3 Drivers attribution agreement

The total accuracy of driver attribution was 96\% using a 0.95 level of confidence (Table 3). *Pasture conversion* exhibited the lowest omission (2\%) and commission errors (1\%) and the highest discrimination from *non-stand replacing* and *agriculture conversion*. *Non-stand replacing disturbance* had 4\% and 14\% omission and commission errors, respectively. The worst agreement was found in the *agriculture conversion* class, which exhibits the highest omission (44\%) error and commission of 9\%. This high error is likely due to small sample size.

A total of 10 patches (5\%) were labeled as *unclassified* because of their low confidence attribution. The total accuracy when measured using area-weighted errors was 98\%±0.008 (Table 4), which indicates a slight reduction in error compared to
patch-level accuracies. Commission’s errors are reduced for agriculture conversion and non-stand replacing disturbance classes whereas omission errors remained almost equal.

Table 4. Confusion matrix at patch level.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Conversion to pasture</th>
<th>Conversion to agriculture</th>
<th>Non-stand replacing disturbance</th>
<th>User’s Accuracy</th>
<th>Commission error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map</td>
<td>141</td>
<td>1</td>
<td>99%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1</td>
<td>91%</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>45</td>
<td>86%</td>
<td>14%</td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>98%</td>
<td>66%</td>
<td>96%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omission error</td>
<td>2%</td>
<td>44%</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Area-weighted error matrix for drivers.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Conversion to pasture</th>
<th>Conversion to agriculture</th>
<th>Non-stand replacing disturbance</th>
<th>User’s Accuracy</th>
<th>Commission error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map</td>
<td>0.851</td>
<td>0.018</td>
<td>0.004</td>
<td>99%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>0.004</td>
<td>0.007</td>
<td>96%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.006</td>
<td>0.009</td>
<td>0.109</td>
<td>88%</td>
<td>12%</td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>99%</td>
<td>67%</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Omission error</td>
<td>1%</td>
<td>43%</td>
<td>5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.4.4 Driver dynamics
Disturbances detected using BFAST Monitor showed a linear increase from 1999 through 2007, except for 2005-2006; after 2007, disturbances had a more inconsistent temporal distribution. The greatest amount of disturbance occurred in 2007 (Figure 15), the result of a significant expansion of pasture (40%) for cattle grazing (Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales, 2016). Of note, the increase in disturbed area in 2007 is not a product of increased Landsat image availability as other study years had the same number or more images. Spatially, disturbances were most prominent over lowlands and along the margins of rivers; at higher elevations, the presence of disturbance was much significantly decreased.

The dominant driver throughout the study period was conversion to pasture, which accounted for 9901±72ha, well in line with CLC-calculated area of 10827ha. The second most dominant driver was non-stand replacing disturbance. Although there is no precedent for documenting such disturbances in Picachos, we estimated an area of 1327±92ha that suggests a progressive degradation (i.e., selective logging) towards a conversion in land cover. Conversion to agriculture contributed 323±61ha, which was in good agreement with CLC estimates of 380ha. Finally, 1467ha corresponding to 12.7% of the total disturbed area could not be confidently attributed and was left as unclassified.
Conversion to pasture was found over the margins of Templado, Plantanillo and Yulo rivers while conversion to agriculture and unclassified patches were mainly positioned over Yulo and less over Platanillo river. Both drivers are near Picachos’ official border, which likely reflects the influence of nearby peasant reserves. Non-stand replacing disturbance was more dispersed over the entire study area and even evident at higher elevations. Notably, conversion to pasture and non-stand replacing disturbance are commonly in close proximity to each other, suggesting a gradual land conversion process (Figure 16). Stable forest, on the other hand, was commonly found at higher elevations (>700m) likely due to more complex topography and limited physical accessibility.

Figure 16. Spatial distribution of classes from 1999-2015.

3.5 Discussion

Using a combined analysis from different topographic, pattern, trend and spectral metrics derived from Landsat TM, ETM+ and OLI sensors during the period 1996-2015 national park, we detected, characterized and attributed changes on Picachos foothills. The presence of disturbances based on NDMI-BFAST indicate a good
agreement for detecting both gradual and abrupt disturbances (Table 4 and 5). The total accuracy for detection of disturbances and stable forest was high, with commission errors smaller (4%) whereas omission error for change class was 21%. The higher rate of omission can be explained by a lower density of available images for this study area (mean value 51%) in comparison with other studies in tropical countries such as: Peru, Brazil, Ethiopia and Vietnam (Devries & Verbesselt, 2015; Schultz et al., 2016).

3.5.1 Method effectiveness

NDMI was successfully applied to detect disturbances and characterize post-disturbance conditions. NDMI leverages the sensitivity of SWIR1, which was especially useful for detecting conversion to pasture areas with high soil reflectance exposure. After a subtle reduction of canopy cover, NDMI was also able to detect regrowth faster than other vegetation indices, which is particularly useful for characterizing non-stand replacing disturbance (Devries & Verbesselt, 2015). NIR, on the other hand, is related to biomass increase and effective at monitoring conversion to agriculture.

Since different drivers happen at different temporal rates the most important metrics for characterization of drivers were the mean and the standard deviation NDMI post-change, NDVI-trend, change magnitude and site variables such as elevation. Pattern metrics such as patch area or perimeter had little contribution to the classification likely because disturbance patch size was largely similar independent of the driver, and larger patches affected by other external agents (i.e. wildfires) are not common in the study area (Rodríguez et al., 2013). The relative importance of these metrics would likely vary if we included other potential drivers of change or expand the analysis over the rest of the Andes–Amazonia transition belt.

Using the second highest voted class, we minimized the assignation of objects towards unreliable drivers, which allowed us to achieve a better mapping confidence and attribution performance. A confidence value per patch is a quantitative indicator of potential confusion between two alternative drivers and provides an opportunity for future sampling design (Mitchell et al., 2008). During this approach, unclassified
areas occupied 12.7% of the area. This information is useful for future field campaigns as well as for testing other hypothesis and datasets to reveal unknown causes of disturbances.

The findings presented here can complement the spatio-temporal inventory of Picachos protected area. In Colombia, the generation of land-cover maps in remote places is carried on every 5 years and it’s still non-validated. This analysis provides annual and disaggregated information from 1999-2015 about the most common drivers of change using all available Landsat imagery and reference datasets. Although we worked merely over Picachos foothills, the final objective is to apply this procedure over the Andes-Amazon transition belt, which includes the protected areas of Tinigua and Macarena.

3.5.2 Method limitation

While we have detected, characterized and attributed disturbances using a set of metrics with a relatively high agreement, there are some limitations for improvement in future research. First, using BFAST in one-year consecutive periods, we were unable to identify multiple disturbances at a given pixel over the time series. Second, the training data used in building the RF classifier was not probabilistically sampled but rather driven by the availability of reference datasets (e.g., CLC, RapidEye and high-resolution images). Despite limitations in our sampling approach, the out-of-bag sample provided by the RF model provided an alternative robust measure of model accuracy that identified an acceptable error rate of 13%. Finally, pre- and post-disturbance spectral values were calculated over all available dates in the time-series, which can introduce possible outliers and are prone to be less effective for the years in the beginning and end of the time-series.

While conversion to pasture and non-stand replacing disturbance were successfully discriminated, conversion to agriculture was commonly confused with non-stand replacing disturbance. This confusion was likely due to similar pre- and post-change spectral conditions in that crop phenology can produce comparable values with subtle change in the short- and long-term (Figure 12). Making a clear distinction between both classes is difficult since agriculture has a relatively low
coverage area and is generally surrounding by other land covers, which can introduce a mixed pixel effect, reducing the potential for accurate classification. Moreover, despite non-stand replacing disturbance often preceding deforestation driven by subsistence agriculture (Devries et al., 2016; Lambin, 1999) or pasture expansion (Fernández et al., 2015; Walker et al., 2009), we were unable to identify causal linkages between specific drivers. In future research, engaging local community knowledge could be relevant for establishing linkages between drivers and across metrics, especially given the remote localization of our study area (Davies, Everard, & Horswell, 2016; Devries et al., 2016; Miranda, Corral, Blackman, Asner, & Lima, 2016).

Other potential drivers such as wildfires and road expansion were not considered in this study. First, wildfires are uncommon in Picachos (Rodríguez et al., 2013) and available fire datasets have a coarse spatial resolution which would be impractical to monitor at patch level. In the case of roads, all of them are unpaved and their small size means they could not be detected at the Landsat scale of observation (Stewart, Wulder, McDermid, & Nelson, 2009).

3.6 Conclusions

We integrated pattern, spectral, topographic, and trend metrics derived from BFAST Monitor and the Landsat archive to detect disturbances and identify related drivers in the Picachos foothills. At patch and area levels, the discrimination and attribution was more accurate for conversion to pasture and non-stand replacing disturbance than for conversion to agriculture with spectral and trend metrics having the most influence on the driver identification. We have confirmed that conversion to pasture is the main cause of deforestation throughout the study period and over Picachos. The temporal and spatial information provided by this methodology can be used to improve the monitoring of disturbances in protected landscapes under threat from anthropogenic drivers.
3.7 References


Cortes Alonso, & Sergio, D. (2012). *Assessing the ground truth drivers of land cover and land use changes at a local scale in Cundinamarca and Tolima Departments in Colombia*. University of Southampton.


Colombia; p. 118.


Chapter 4. General Conclusions

This research demonstrates the opportunities to assess ecological changes through multi-source remote sensing, sub-annual change detection algorithms and machine learning techniques in regions with persistent cloud cover. Using Landsat and MODIS-based MAIAC archives, I identified short- and long-term forest disturbances as well as the potential drivers of change with a relatively high spatio-temporal accuracy from 1999-2015 in Picachos Natural Park.

Ecological indicators such as greening and browning were mapped using dense MAIAC data within Picachos and its surrounding 10km buffered area. I found a hotspot of vegetation decline in the park’s southeastern region that likely results from population incursion and land use conversion. This result was corroborated using Landsat imagery, where disturbed area in the park was 12,642 ha (±1440 ha) corresponding 4.3% of Picachos’ total area. At higher elevations (>2000m) the lack of MAIAC and Landsat observations reduced my ability to detect disturbance. With a hotspot of vegetation decline detected and using a set of pattern, spectral, trend, and topographic metrics, I classified drivers of change. Pasture expansion for cattle grazing is the main cause of change within Picachos’ foothills, followed by forest thinning and agricultural development.

The results of this study can be used by Protected Areas managers to expand document localized disturbances despite relative inaccessibility. More broadly, the promising results of this study should encourage the remote sensing community to develop more robust frameworks using dense time-series metrics to analyze the drivers and consequences of human-induced changes, especially in regions with low imager availability or consistent cloud cover. Valuable future research in the Picachos region would focus on the impact of armed conflict on forest changes, internally displaced people (IDPs), along with illegal land appropriation. The effect of forced displacement and population resettlement on potential forest recovery and degradation in Picachos would be a timely and relevant study, especially in the context of the Colombian post-conflict period. Expanded knowledge of particular pressures on the landscape and recognition of its complex socio-economic dynamics
will enrich the implementation of sustainable strategies in this region that support people and forests alike.

**Bibliography**


Cortes Alonso, & Sergio, D. (2012). *Assessing the ground truth drivers of land cover and land use changes at a local scale in Cundinamarca and Tolima Departments in Colombia*. University of Southampton.


Instituto de Hidrología Meteorología y Estudios Ambientales de Colombia (IDEAM). (2010).
Leyenda nacional de coberturas de la tierra. Metodología CORINE Land Cover Adaptada para Colombia Escala 1:100000. Bogotá, Colombia.


Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and


### Table 6. Total number of Landsat images used for disturbance detection.

<table>
<thead>
<tr>
<th>Year</th>
<th>Landsat 5</th>
<th>Landsat 7</th>
<th>Landsat 8</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1997</td>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>1998</td>
<td>8</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1999</td>
<td>7</td>
<td>2</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>7</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>2001</td>
<td>6</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>2002</td>
<td>6</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>2003</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2004</td>
<td>8</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>2005</td>
<td>6</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>2006</td>
<td>5</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>2007</td>
<td>1</td>
<td>8</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>2008</td>
<td>7</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>2009</td>
<td>10</td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>2010</td>
<td>1</td>
<td>9</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>2011</td>
<td>1</td>
<td>7</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>2012</td>
<td>10</td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>2013</td>
<td>9</td>
<td>1</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>2014</td>
<td>4</td>
<td>9</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>2015</td>
<td>8</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>105</td>
<td>18</td>
<td>149</td>
</tr>
</tbody>
</table>
Figure 17. Percentage of cloud-free Landsat observation in the period 1996 –2015.
Figure 18. Selection of magnitude threshold ($P(disturbances)=0.5$) using Binomial Logistic Regression. The grey area is 95% confident interval. Training sites (blue points) were selected based on visual interpretation where 0=non-disturbed and 1=disturbed.