UNDERSTANDING PATTERNS OF LAND-COVER AND LAND-USE CHANGE IN THE TRIFINIO REGION OF CENTRAL AMERICA

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctorate of Philosophy
with a
Major in Natural Resources
in the
College of Graduate Studies
University of Idaho
by

Peter Schlesinger

Major Professors: Lee A. Vierling, Ph.D.; Miguel Cifuentes-Jara, Ph.D.
Committee Members: Kelly W. Jones, Ph.D.; Pablo Imbach-Bartol, Ph.D.;
Jan U.H. Eitel, Ph.D.
Department Administrator: Lee A. Vierling, Ph.D.

December 2017

Authorization to Submit Dissertation

This dissertation of Peter Schlesinger, submitted for the Degree of Doctorate of Philosophy with a Major in Natural Resources and titled "**Understanding Patterns of Land-Cover and Land-Use Change in the Trifinio Region of Central America,**" has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professors: _____ Date: _____
Lee A. Vierling, Ph.D.

_____ Date: _____
Miguel Cifuentes-Jara, Ph.D.

Committee Members: _____ Date: _____
Kelly W. Jones, Ph.D.

_____ Date: _____
Pablo Imbach-Bartol, Ph.D.

_____ Date: _____
Jan U.H. Eitel, Ph.D.

Department
Administrator: _____ Date: _____
Lee A. Vierling, Ph.D.

Abstract

Land-cover and land-use studies in Asia, Australia, and Canada began in the early 1970s with the advent of the first Landsat sensors. However, the satellite receiving stations didn't have extensive coverage until the mid-1980s. With the advent of the new century, regional onsite researchers began noting forest regrowth signs in Central America. However, few investigations took place in the Trifinio Region, a transboundary region at the join of El Salvador, Guatemala and Honduras. Most were considerably beyond, in the Yucatan, the Guatemalan highland, central El Salvador, and the Peten.

In Chapter 1, we examined secondary maps to understand the region, its human and natural resources and changes. We sought to create new modeling, aided by data from participating agricultural development partners. We considered land use data to learn about correlated change drivers and potential transition causes and studied transitions inside and outside the region's most diverse protected cloud forest park. We found population density correlated with severe deforestation, found a relation between transport and deforestation, and found potential drivers of deforestation to be much the same as those affecting the wider region in Central America and tropical zones.

A bold assumption also underpinning Chapter 2 was that we ought to be able to sense, measure, and compare common results in terms of vegetative responses and differences across the region's countries, counties, and protected areas (PA). We developed large databases, sampling the territories of the protected and non-protected landscapes of Trifinio Region to explore forest transitions that we could not measure using classical remote sensing classification methods. We prepared percent greenness from NDVI and created novel methods to describe and understand temporal transitions both inside and about 20km outside of the Trifinio Region because there were no tools ready for us to use our short time series data readily, though we found existing gene expression tools useful. We hunted for a sign of forest regrowth or resurgence connected with the regional protected areas, agriculture, pastoral activity, and agroforestry, to corroborate reports from regional studies in the past two decades. In follow-on work in Chapter 3, we calculated pattern and texture measures for an area of central Trifinio Region using different types of relatively new satellite sensors (both optical and radar) with a view to being able to identify differences in shade agroforestry from

native forest cover. We found relationships between the two of these measures, shade coffee production, and forest cover. We extracted training sites using a high-resolution sensor to classify these areas using machine learning classifiers; we compared our efforts accuracy estimates with those of other researchers, some of whom had tried some of these classifiers before us. Few experiences have been reported studying rural agriculture with these sensors and techniques.

Acknowledgements

I would like to thank my major advisors Lee Vierling and Miguel Cifuentes for the time they spent working with me on this endeavor, for each meeting, formal or informal, for every Skype, every office get-together, planned or unplanned, for every bit of time they gave me to put me on the path to completion. Thanks so much above all for your patience and understanding. I thank my committee members Kelly Jones, Pablo Imbach, and Jan Eitel for all the comments, time, ideas, and guidance you gave me at the key moments. Thanks for being available to me! I thank you all for sharing your ideas, knowledge, and experience. Thanks for all the guidance and for your friendship.

I want to thank the University of Idaho and the National Aeronautics and Space Administration for their financial support of our project in the Trifinio Region of Central America that created the funds to permit me to take part in the Joint Doctoral Program between the University of Idaho and the Tropical Agricultural Research and Higher Education Center (CATIE). The University of Idaho initially provided two years of funding and project support monies for travel, onsite accommodations, project support, and meetings attendance through support by the Land-Cover and Land Use-Change Program through a grant to the College of Natural Resources. In my last semester, the University of Idaho granted me funding support through the Graduate Fellowship Program. Additional support came through the provision of satellite images by the U.S. Fish and Wildlife Service through Northern Arizona University, thanks to the assistance of former CATIE/UIdaho grad Dr. Steve Sesnie.

On the CATIE staff, I'd like to thank Director Muhammed Ibrahim, Deputy Director, Mario Piedra, Dean Isabel Gutierrez, Cynthia Mora, Aranjid Valverde, Marta Gonzalez, Diana Fuentes, Arturo, Jeanette Solano, Kenneth Royo, and all the staff of the guards and maintenance department, Juan Carlos Zamora, Sergio Vilchez, and Alejandra Ospina. Christian Brenes, Fredy Argotty, Alan Coto, Allan Guerero, and Lenin Corales for lending me his chair.

Also in Moscow, a big thanks to UI staff members Cheri Cole, Lynne Kittner, Lana Unger, Katherine Clancy, Bruce Godfrey, and Jory Shelton. Thanks so much to Sanford Eigenbrode and Nilsa Bosque-Pérez for their work with the UIdaho JDP and their assistance.

I want to thank all my professors, colleagues, classmates who were a part of my journey, especially, my colleague and friend, Carlos Munoz and his wife, Shirley Murillo, and especially young Hector Munoz for reminding me why I had embarked on this journey in the first place.

Thanks to all my new friends in Guatemala, Victor Hugo Ramos, Ingrid Hausinger, Kenset Rosales, Abner Jimenez, and in El Salvador, particularly, CATIE staff member, Ana Cristela Gutierrez of San Salvador.

Thanks to my many new CATIE student friends, Fredy Duque and Tati, Carmen, Dulce, Dulcinea, Laura, Pablo Chacon, Roberto, Sebastian and family, Gerardo and family, Grace and Diego, Simon and Lizbeth, and Naylea, Waldir and Ana, and Ayeni, Carlos and Magdalena, and Óscar, Aline, Zayra, Claudia, Marilyn Manlow, Allan Coto, and Natalia Ruiz, who was a major help. And, thank you to the members of the CATIE Consejo Estudantil 2016, the 2016 CATIE Student Council.

Dedication

During the time of this research and writing,
I lived in Moscow, Idaho, and Turrialba, Costa Rica.
Gracias to my Costa Rican friends who made life more fun:
Ricardo Vega and Bernardo of Turrialba Rent-A-Car,
Dandy “Mynor” Carranza Quesada the Pipa Man, and Rosalba Gil.

A special thank you to my mom and dad,
Pat and Tom Schlesinger of New Hampton, New Hampshire,
without whom this would not have been possible;
my sister, Annie, who helped me to dot the i’s and cross the t’s;
my wife, Jennifer Bendezu, for taking care of me and putting up with my angst;
and my son, Sam, and daughters Andrea, Alanis, and Annie.

Table of Contents

List of Figures

List of Tables

List of Equations

CHAPTER 1: THE TRIFINIO REGION: A CASE STUDY OF TRANSBOUNDARY FOREST CHANGE IN CENTRAL AMERICA

"The Trifinio Region: a case study of transboundary forest change in Central America." *in Journal of Land Use Science,* 12(1), 36–54.

Abstract

Strategic planning to increase forest cover in Central American transboundary areas of the Mesoamerican Biological Corridor requires understanding land-cover and land-use change trends and drivers. We estimated forest cover change from remotely-sensed land-cover and land-use classifications from 1986, 2001, and 2010, in the tri-national Trifinio Region, bordering El Salvador, Guatemala, and Honduras. Our analysis spanned sub-national, national, regional, and protected area borders. We determined correlations with direct drivers of deforestation, developing a multilevel linear regression model. Higher population density was significantly correlated with deforestation; coffee, agroforestry, and pasture replaced forests. The tri-national park retained forests compared to neighbouring areas, but additionality requires more research. The literature on drivers suggests similar processes and factors in other tropical regions. Forest cover governance efficacy is highly variable. Results indicate relationship between governance and forest cover; though more comprehensive understanding of this complex region is needed to determine their causality.

Keywords: Coffee, deforestation, land use, protected area, transboundary, Trifinio

Resumen

Las áreas transfronterizas centroamericanas del Corredor Biológico Mesoamericano requieren comprender las tendencias y los impulsores del cambio de uso de la tierra y la tierra. Estimamos la cubierta forestal a partir de las clasificaciones de uso del suelo y de la cobertura terrestre detectadas remotamente desde 1986, 2001 y 2010, en la región tri-nacional Trifinio, que limita con El Salvador, Guatemala y Honduras. Nuestro análisis abarcó fronteras de áreas subnacionales, nacionales, regionales y áreas protegidas. Determinamos las correlaciones con los impulsores directos de la deforestación, desarrollando un modelo de regresión lineal multinivel. Una mayor densidad de población se correlacionó significativamente con la

deforestación; el café, la agrosilvicultura y el pasto reemplazaron a los bosques. El parque tri-nacional preservo los bosques en comparación con las áreas vecinas, pero la adicionalidad requiere más investigación. La literatura sobre conductores sugiere procesos similares y factores en otras regiones tropicales. La gobernanza de la cubierta forestal tiene una variable muy alta. Los resultados indican una relación entre gobernanza y cobertura forestal; aunque un entendimiento más comprensivo de esta compleja región es necesario para determinar su causalidad.

Palabras clave: Café, deforestación, uso del suelo, área protegida, Transfronterizo, Trifinio

Introduction

Effective governance of land-cover in transboundary PAs of Central America and their connecting corridors is vital for regional biodiversity conservation (Sandwith, 2001). However, nearly 4000 km2 of forest in the region is used annually for farming (Graham, 2002; Inter Press Service, 2002; Food and Agricultural Organization [FAO], 2010d), undermining regional biological corridors (Mendoza et al., 2013), as well as the efficacy of protected area governance (Stevens, 2013). Barriers to effective transboundary governance in forest conservation can include differences in access to monitoring technologies and the ability to work together with common goals regardless of administrative boundaries and national priorities (Sandwith, 2001). Transboundary governance efficacy should be measurable, when decoupled from municipal and national administration, with implementation of unilateral governance policies (Zbicz, 1999). It is unclear how much these governance approaches affect land-cover and land-use change (LCLUC) and which of the underlying drivers of this change are most important.

Two examples of transboundary PAs in the Mesoamerican Biological Corridor are the Maya Biosphere Reserve (MBR) (Belize-Guatemala-Mexico) and La Amistad (Costa Rica-Panama). The MBR is under severe LCLUC pressure due to subsistence agriculture and ranching combined with illicit trafficking; illegal logging and other incursions have taken a significant toll on forests (Carr, 2008a; Choi, 2008). Governance and LCLUC monitoring have been a joint effort by government, military, and nonprofit organizations using climate

mitigation and forest certification as conservation tools. At La Amistad, one of the oldest and largest transboundary PAs, governance mechanisms have been ineffective at countering LCLUC (Oestreicher et al., 2009), requiring dynamic surveillance and strong command and control measures. In neither case was true transboundary management initiated; the PAs extend across borders, but their territories are managed separately within each country as international adjoining PAs, as described by Zcibz (2001).

A third transboundary protected area, the Montecristo Trifinio Protected Area (MTPA), resides within the Trifinio Region situated between El Salvador, Guatemala, and Honduras (Organization of American States [OAS], 1993; Artiga, 2003). The MTPA has been managed as a tri-national conservation area since 1987; in 1997, 221 km2 were set aside as a tri-national protected area, spanning adjoining boundaries (CTPT, 1997). In 2011, it became a core site of the Trifinio-Fraternidad Biosphere Reserve under UNESCO-MAB Programme and the first tri-national biosphere reserve in Central America. Trifinio's governance and LCLUC drivers pose challenges for resource management, because like many regional border areas, Trifinio is poor, marginalized, and isolated from national political centers (Girot, 1997). Yet the Trifinio region is also rich in biodiversity, water resources, and culture, with a full range of optimal elevations for coffee agroforestry (400-1600m). While strategic planning agencies have promoted rural development and water resource protection in Trifinio for more than two decades, little peer-reviewed regional scientific work has been published on this area, with the exception of a few studies focused on species habitat and biodiversity (e.g. Komar, 2002; Greenbaum & Komar, 2005).

In this paper, we advance the knowledge base on transboundary governance by investigating LCLUC at different administrative levels in the Trifinio Region. Our research objectives are: 1) to assess the region's major land-cover and land-use changes and differences across governance boundaries, including the MTPA, and 2) to identify direct drivers of deforestation affecting this unique tri-national transboundary area. Related to objective one, digital maps estimating forest cover and major land uses were created to support regional rural development planning for three time periods: 1986, 2001, and 2010 (Water Center for the Humid Tropics of Latin America and Caribbean [CATHALAC], 2011). We used these unpublished data to quantify LCLUC within transboundary and protected area governance units. A key assumption is that the trends identified in our assessment occurred in

response to transboundary management and protective strategies, which are measurable, common, and constant across the region. Specifically, we assess: What kinds and amounts of change are apparent? Is there a relationship between LCLUC and the MTPA? Is LCLUC more responsive to certain levels of nested governance than others?

The integration of theories and methods from socioeconomic, remote sensing, and ecological sciences has been used to identify key indirect (underlying) and direct (proximate) drivers of land change (Angelsen & Kaimowitz, 1999; Geist & Lambin, 2002; Lambin & Geist, 2006; Turner, Lambin, & Reenberg, 2007; Lambin et al., 2014). In this paper we develop a multilevel linear regression model to assess the levels of correlation among direct drivers of change, different administrative levels and deforestation. We also highlight the potential indirect drivers of LCLUC in Trifinio through secondary literature review. Specific research questions assessed related to our second objective include: What are the key events and drivers of LCLUC in the Trifinio Region? Are these events and drivers the same as found elsewhere in land use literature? Is LCLUC correlated with certain levels of nested governance more than others?

Methods

Study Area

Establishment of Trifinio Region

The Trifinio Region (Figure 1.1), comprising 7,500 km^2 in El Salvador (15%), Guatemala (45%) and Honduras (40%), is an outcome of peace accords signed August 7, 1987 to end 40 years of war. The war had tremendously negative implications for economic development, institutional reform, landlessness, and natural resource uses that are still apparent today. One tri-national partnership that arose out of this period spurred the Comisión Trinacional del Plan Trifinio, a group of member country vice presidents, advised by municipal and civil society organizations (Miranda, Slowing Umaña, & Raudales, 2010). This Trifinio commission manages and promotes a mission of regional conservation and sustainable use of natural and water resources.

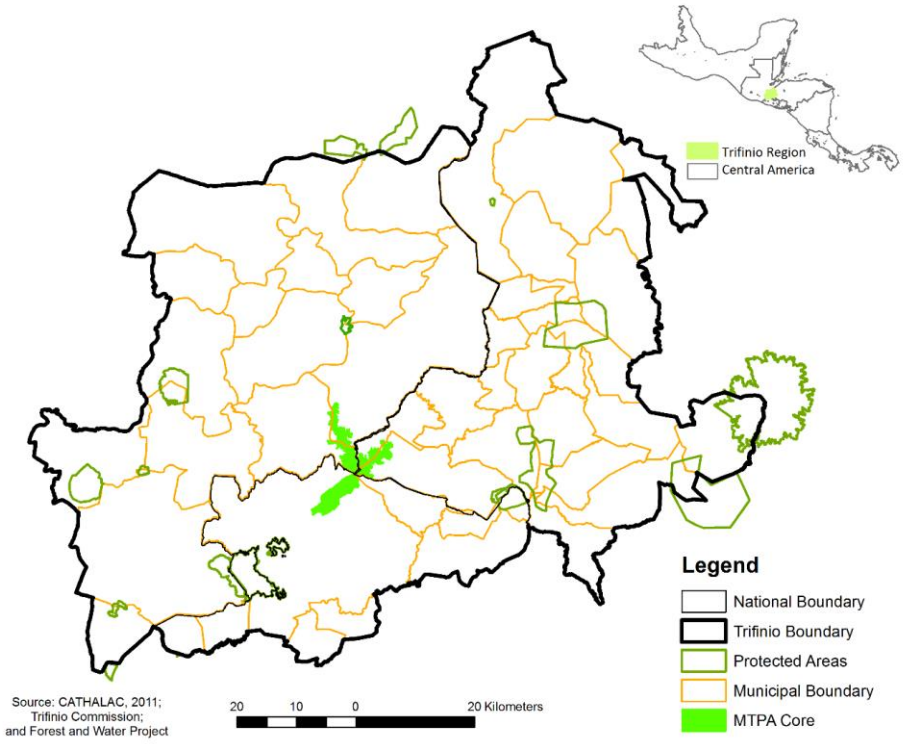


Figure 1.1. Map of the Trifinio Region location in Central America, municipalities by country, the core of the Montecristo Transboundary Protected Area (MTPA), regional PAs, and the nested protective governance structure.

Socioeconomics

Regional population estimates show that Trifinio has grown from 572,000 to nearly 900,000 during 1987-2011 (German Society for International Cooperation (GIZ), 2011). These estimates are based in part on projections, in the absence of recent censuses for Honduras (2001) or Guatemala (2002). Marginalization estimates report 87% in poverty and 53% in extreme poverty in the region; however, these too are out-dated (Artiga, 2002). Nonetheless, extreme poverty inhibits farmers, who maintain an average of 1.4 ha of land, from becoming development actors in their own right (Castaneda, 2009; Cherrett, personal communication 2014).

Governance

Trifinio is managed using a series of spatially and institutionally-nested regions intended to promote cross-border cooperation and integration. Nested, multi-level governance is common in Central America (Celata, Coletti, & Sanna, 2013), but Trifinio's

implementation is unique in the transboundary schema. It consists of three distinct nested regions – a regional outer boundary, a central area for sustainable development, and a highly protected core – and promotes international inter-municipal integration (MARN, 2010). Celata et al. (2013) categorize this as an example of a new type of territorialization (Figure 1.1).

There are 45 municipalities in Trifinio (22 in Honduras; 15 in Guatemala; and eight in El Salvador) (Figure 1.1), including several hundred small villages and six cities. Multilateral aid supports forest management and reforestation, as well as basic services in these municipalities (Global Environment Facility [GEF], 2006). Between 1998 and 2006, groups of 3-5 municipalities further organized into 'Mancomunidades'; these governing entities are meant to facilitate cohesion and cooperation and foster economies of scale for productive activities and sustainable land use.

The MTPA in Trifinio has a core zone straddling the three country borders, surrounded by a buffer and multiple use zones (Figure 1.2). Its biological importance is for mountain pine-oak and Pacific dry forests, highly-threatened ecoregions of Central America, and a key element of regional conservation efforts over the last 20 years (Dinerstein et al., 1995). It is also the source of three major river systems: the Motagua, Ulua, and Lempa. The Lempa, one of the largest watersheds in Central America, serves 4.7 million downstream users alone (InterAmerican Development Bank [IDB], 2010). Cooperation and integrated management of the riparian area is critical for downstream populations (Artiga, 2002; United Nations Environment Program [UNEP] & Permanent Commission for the South Pacific [CPPS], 2006; Carius, 2007).

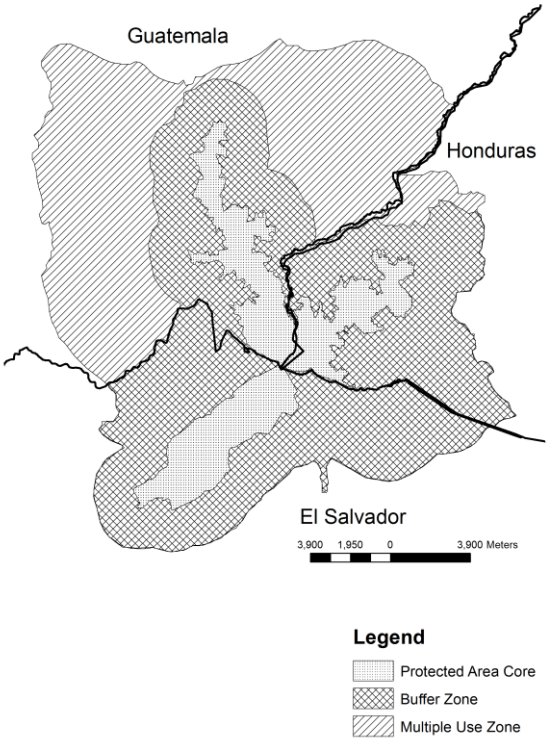


Figure 1.2. Zones of the Montecristo Trifinio Protected Area.

Data

Our research draws on site visits to Trifinio to collect secondary remote sensing data, geospatial data, interview notes, documents, and photos taken with corresponding GPS points that were acquired to help better understand the land-cover and the map classification in August and November 2013. The Water Center for the Humid Tropics of Latin America and Caribbean (CATHALAC) led a group of 25 regional agency and university participants who mapped Trifinio's natural resources using satellite images through an InterAmerican Development Bank/Global Environment Facility-funded project, 'Integrated Management of the MTPA' (CATHALAC, 2011). Their transboundary land-cover classifications were produced for three epochs – 1986, 2001, and 2010 – with support from the GIZ Forests and Water project. Our study compared and quantified rates of change that we estimated using three land-cover and land use geospatial datasets (i.e. shapefiles) produced by the CATHALAC-led regional project. These CATHALAC shapefiles were the spatial final result of a careful classification of least-cloud Landsat 5 Thematic Mapper images following standard remote sensing methods (using ENVI v. 4.5, ERDAS Imagine v. 9.2, and ArcGIS v.

9.3) (see Appendix 1 for more details). We quantified LCLUC using the mapped classifications, as described in as described in the Analysis section below.

We acquired administrative boundary and associated population data to facilitate regression model development. Road density was prepared from the global gROADSv1 dataset (CIESIN, 2013) dividing summed road segment length for each municipality (km) by municipal area (km^2). Population density was prepared similarly, dividing population (CTPT, 2011) by municipal area in km^2. Elevation and slope were estimated using 30m data from the Shuttle Radar Topography Mission (Farr et al., 2007). We derived novel forest loss rates using the CATHALAC land-cover and land use epoch shapefiles (CATHALAC, 2011).

Types and differences in land-cover and land-use changes

We used the CATHALAC land-cover classifications (Table A1.1) to create deforestation layers corresponding only to areas of forest loss during each of the three epochs occurring from 1986-2001, 2001-2010, and 1986-2010. These and all other raster and vector input data were re-projected to Universal Transverse Mercator (UTM) Zone 16-North using map datum WGS1984 via established geographic analysis tools – IDRISI Selva® version 17.02 (Eastman, 1989) and ArcGIS® version 10.1 (ESRI, 2012). We used the IDRISI Selva CROSSTAB function to estimate all transitioning classes via cross-classification. We calculated forest cover annual rates of change using Puyravaud's formula (2003, p. 594),

$$r = \frac{1}{t2-t1} \ln \frac{A_2}{A_1} \qquad \text{(Equation 1.1)}$$

where t_1, t_2 = Time 1 and Time 2; and A_1, A_2 = Forested Area in Time 1 and Time 2. This method represents a practical, conservative method of calculating annual rates of LCLUC (FAO, 1995).

Our objective was to estimate pan-Trifinio forest cover changes, with comparisons across national boundaries, followed by a narrower focus to consider changes in 250 m-wide Euclidean distance-based buffers within and up to 4 km-distant from MTPA's core zone. Buffers were normalized by their area, and a 300-m mask in the central part of the MTPA eliminated imperfections of adjoining national boundaries (see Figure 1.2). Our approach followed that of Joppa, Loarie, & Pimm (2008), who examined polygons inside and outside of the World Database of PAs (WDPA) to examine percent deforestation and fragmentation, as

well as Nepstad et al. (2006), who quantified fires and deforestation in and outside of parks and reserves in the Brazilian Amazon. Comparisons of forest cover area for each buffer were summarized in 1 km increments for countries and municipalities over the study period (Figure 1.3). The total area of our analysis was reduced to about 7,300 km^2 (from 7,500 km^2) due to combined insignificant classes (e.g. lava, mining, unclassified, etc.).

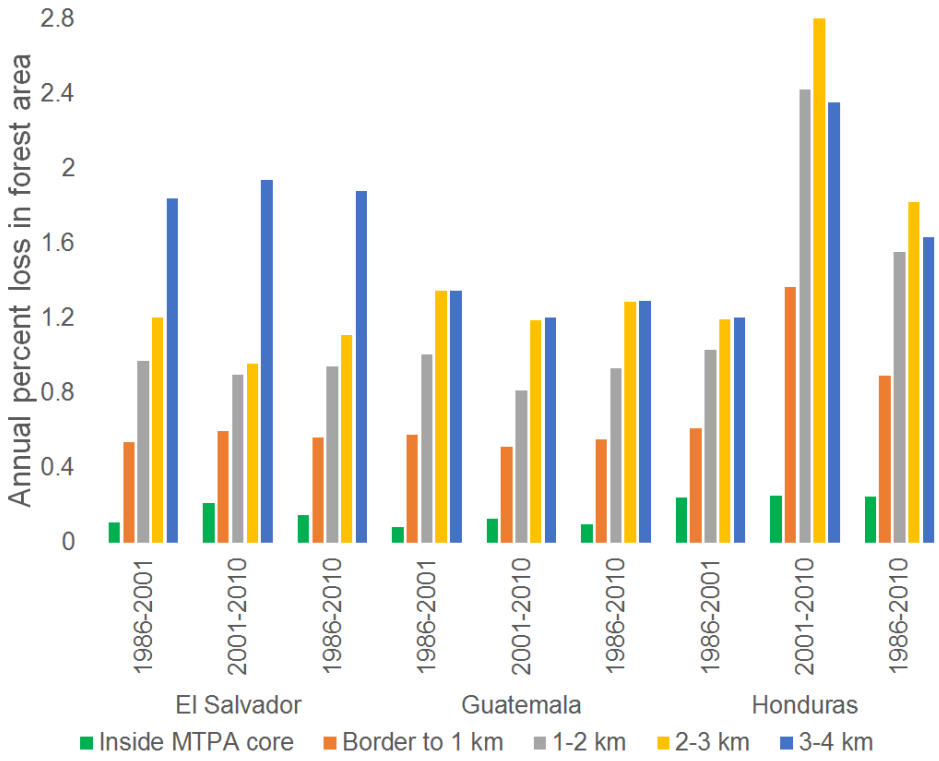


Figure 1.3. MTPA buffer analysis showing the tapering off of deforestation from the protected area core, based on CATHALAC (2011) data.

Direct drivers of deforestation

We focused on changes in deforestation and the associated drivers, utilizing secondary data sets, field interviews, and mapping evaluation. The focus on deforestation, whether driven by ecological marginalization of farmers or by public or private investments for frontier development, is justified as an indicator of the potential economic use of land in this transboundary area (Geist & Lambin, 2002). The effects of deforestation at fine spatial scales in Trifinio is likely to follow similar patterns as in other tropical regions, with increases in land surface temperature and reduction in evapotranspiration with forest conversion to pasture and agriculture (Lawrence & Vandecar, 2015).

We developed a multilevel linear regression model to determine correlations between direct drivers, municipality- and country-boundaries, and forest cover loss for the Trifinio Region. Multilevel models are advantageous for land-cover change analysis when data have a nested structure (e.g., Pan & Bilsborrow, 2005, Vance & Iovanna, 2006). Multilevel models relax the assumption of independence between observations by decomposing the error term into hierarchical components – in this study municipalities are nested within countries – and then imposing a structure on the variance and covariance of these terms. This controls for correlations across municipalities within the same country and has emerged as an alternative strategy in correcting for spatial correlation when the correlation has a nested structure (Anselin, 2002). Maximum restricted likelihood was used to estimate the multilevel model.

The multilevel model was specified for two levels: a level-one municipality-level effect and a level-two country-level effect. The level-one model, where *i* is a municipality, *j* a country identifier, and *t* time, can be expressed as:

$$Y_{ijt} = \varphi_{0j} + \tau F_{it} + \beta X_{ijt} + \vartheta T_t + d_{ij} + \varepsilon_{ijt}, \qquad \text{(Equation 1.2)}$$

where Y_{ijt} is the amount of forest cover loss in municipality *i* nested in country *j* at time *t*; φ_{0j} is the country-specific effect for all municipalities within the same country; F_{ijt} is total forest area; X_{ijt} is a vector of covariates controlling for direct drivers of forest loss; T_t is a vector of time dummy variables capturing time-varying and spatially invariant unobservables; d_{ij} is a municipality-level error term; and ε_{ijt} is residual error. The level-two effects enter Equation 1.2 as:

$$\varphi_{0j} = \delta_{00} + \mu_{0j}. \qquad \text{(Equation 1.3)}$$

μ_{0j} is the country-specific random effect and δ_{00} is the average outcome for the population. Combining these two equations (Equations 2 & 3) gives the full multilevel structure:

$$Y_{ijt} = \delta_{00} + \tau F_{ijt} + \beta X_{ijt} + \vartheta T_t + d_{ij} + \mu_{0j} + \varepsilon_{ijt}. \qquad \text{(Equation 1.4)}$$

Y_{ijt} was derived from the CATHLAC land-cover data, providing two epochs of forest cover change: 1986-2001 and 2001-2010. To ease interpretation of regression output, we converted the 10-year change periods into yearly rates of change by dividing total change by ten. We log-transformed the dependent variable because the rate of forest loss was skewed

and log-transforming it led to a more normal distribution. Total forest area, F_{ijt}, was included to control for the stock of forest resources available at the start of the time period since our dependent variable was the area of forest loss and not the rate of forest loss. Direct drivers, X_{ijt}, included in Equation 1.4 are: total protected area, distance to national capital, road density, population density, elevation, and slope. We log-transformed each of these independent variables to ease interpretation of coefficients; with the log-log model we can look at elasticity (i.e., percent change in X leads to a percent change in Y).

By estimating Equation 4 without covariates (i.e., without F_{ijt} and X_{ijt}) we can calculate the unconditional intraclass correlation coefficient and the proportional reduction in total residual variance for the specifications with covariates (i.e., R^2 for the multilevel model). Rabe-Hesketh & Skrondal (2008, p.58) give the formula (Equation 5) for the intraclass correlation as:

$$\hat{\rho} = \frac{\hat{\psi}}{\hat{\psi}+\hat{\theta}+\hat{\omega}}\,, \qquad \text{(Equation 1.5)}$$

where $\hat{\psi}$ is the estimated variance from the country-level effect, μ_{0j}; $\hat{\theta}$ is the estimated variance from the municipality-level effect d_{ij}; and $\hat{\omega}$ is the residual variance from ε_{ijt}. As written, this formula gives the percent of variation in forest cover loss attributable to the country-level; the amount attributable to municipalities and observations is found by substituting the appropriate variance component into the numerator. Following this notation, the total variance from the null model can be calculated in Equation 6 as:

$$\hat{\eta}_0 = \hat{\psi} + \hat{\theta} + \hat{\omega}\,. \qquad \text{(Equation 1.6)}$$

Letting $\hat{\eta}_1$ represent the total variance from Equation 4 with covariates, the formula for R^2 (Equation 7) is given in Rabe-Hesketh & Skrondal (2008, p.102) as:

$$R^2 = \frac{\hat{\eta}_0-\hat{\eta}_1}{\hat{\eta}_0}\,. \qquad \text{(Equation 1.7)}$$

While countries and municipalities are treated as random effects in Equation 1.4, Rabe-Hesketh & Skrondal (2008, p.77) describe a method using maximum likelihood estimation that can estimate values for these random intercepts. Briefly, this method assumes that the estimated parameter values from Equation 4 are the true values and that the random intercepts are the only unknown parameters to be estimated. The values for each country-level

and municipality-level intercept that maximizes the likelihood of the observed responses of the dependent variable after treating the model parameter estimates as known are then calculated. These estimated values can be interpreted as the country-level and municipality-level influence on unexplained variation in forest cover loss. Thus, it represents the unexplained impact a country or municipality has on forest cover loss, after controlling for the covariates in Equation 1.4.

Results

Types and rates of land-cover and land-use changes

Pan-Trifinio

Our review of the CATHALAC land-cover classification showed that the proportional distribution of forested area in 1986 by country was 9% in El Salvador, 29% in Guatemala, and 62% in Honduras. By 2001, the forested area in El Salvador (8%), Guatemala (27%), and Honduras (65%) had changed slightly, and by 2010, it had further fluctuated again to 9%, 26%, and 65%, respectively. The major changes seen across land uses were coffee, agroforestry, pasture, and urban growth. Agroforestry accounted for mixes of permanent crops, such as maize, beans or coffee, with trees, including plantains, bananas, citrus, or local wood species. No differentiation was made in the land classification for sun-grown and/or shaded coffee.

Across the Trifinio Region, substantial LCLUC occurred from 1986-2010, with a 1.5% annual forest conversion rate (see Equation 1.1). In 1986, forested area was about 3,300 km^2 (about 45% of Trifinio), while coffee (0.1%) was almost non-existent. Losses in forest were mainly driven by on-farm expansions, particularly in coffee, agroforestry, basic grains, pasture and other crops (Table 1.1). The land uses with the highest annual rates of change were coffee (7.2%), pastures (2.5%), and agroforestry (2.3%), representing an increase of nearly 75 km^2 of coffee, 43km^2 of pastures, and 130 km^2 of agroforestry. By 2001, forested area diminished to nearly 2,910 km^2 and by 2010, again to 2,302 km^2. The annual rate of change in forest was highest in 2001-2010, at 2.6% per year, compared with 0.9% per year in 1986-2001. When compared to globally available datasets on forest cover, we found four times greater land-cover change across the Trifinio Region through our analysis of the CATHALAC data than found by the latest global efforts (Hansen et al., 2013).

Table 1.1. Percentage annual rate of change per land use class per country in Trifinio by years of change between epochs.

	El Salvador		Guatemala		Honduras	
LCLUC class	86-01	01-10	86-01	01-10	86-01	01-10
Forests	-1.2	-1.3	-1.3	-3.0	-0.6	-2.6
Coffee	7.3	14.6	8.4	6.0	5.8	9.2
Agroforestry	1.7	1.6	2.2	2.8	2.0	3.1
Basic grains	1.3	1.6	1.6	0.3	1.7	-0.1
Other crops	0.4	1.2	-0.6	6.8	1.9	6.7
Pastures	1.2	1.3	1.6	1.9	2.3	5.1
Bushes	-0.6	-0.9	-0.3	-0.2	0.3	2.1
Urban	5.9	0.9	4.7	2.1	4.5	1.1

Protected area comparison across buffers

The core of the MTPA remained relatively intact (Figure 1.2), while there was significant LCLUC beyond this zone, although this disturbance appears to taper off after the first 3 km distance buffers for El Salvador and within the first 3 km distance buffers for Honduras and Guatemala. Annual rates of change in forest loss increased across the region over time. In El Salvador, the annual rate of change was slightly greater during the 2001-2010 period within the MTPA and within a 1 km buffer from the park; however, rates of change were higher in 1986-2001 for buffers >2 km. In Guatemala, the annual rate of change was higher in 1986-2001 within the MTPA and buffers tested than during 2001-2010. For Honduras, the annual rates of change were more accelerated in the 2001-2010 period for all buffers compared to 1986-2001, but change within the MTPA core zone remained similar across the two time periods. Growth in the urban class was observed in larger and important regional cities (Esquipulas, Metapán, and Ocotepeque), located just beyond the 4 km buffer from the core of the MTPA. This assessment of rates of change in the protected area does not establish a causal relationship between the presence of the MTPA and forest protection (Joppa & Pfaff, 2010; Meyfroidt 2015); it only highlights trends.

Direct drivers of land-cover and land-use change

The multilevel regression structure allowed us to explicitly calculate the municipality- and country-level influence on forest cover loss and their contribution to unexplained variation in the model after controlling for direct drivers of deforestation (Rabe-Hesketh & Skrondal, 2008); this provided some information on the influence that nested governance structures had on LCLUC. Summary statistics of variables included in Equation 4 are found in Appendix 2 in Table A2.1. Variables statistically correlated with forest loss in our regression model included: total forest cover, distance to national capital, road density, population density, elevation and slope (Table 1.2). On average, greater loss occurred in municipalities with more forest cover and people, farther from national capitals, with fewer roads and steep slopes at lower elevations. The amount of variance explained by this econometric model was 79%.

Municipality-level variation accounted for about 40% of total variation in forest loss, and country-level variation accounted for about 16%. This means that more heterogeneity in deforestation patterns exists within a country, i.e., at the sub-national level, than across countries. When we estimated the contribution of nested governance of countries and municipalities to unexplained variation in the regression model, however, coefficients were not statistically significant (Table 2). Figure A2.1 illustrates the direction of unexplained influence of each municipality: a positive sign indicates that a municipality had unexplained variation in forest loss above the average value of forest loss, and a negative sign indicates that a municipality had unexplained variation in forest loss below the average value; while not statistically significant, they show that the general trends vary across and within countries.

Table 1.2. Multilevel regression output of direct drivers of forest cover loss.

Variable	**Mean (Std Dev)**
Amount of forest	0.32***
	(0.02)
Amount of protected land	0.01
	(0.01)
Distance to national capital	0.37***
	(0.10)
Road density	-1.00***
	(0.30)
Population density	0.15**
	(0.05)
Elevation	-0.51***
	(0.12)
Slope	-0.29**
	(0.13)
2001-2010 time dummy	0.33***
	(0.04)
Observations	*90*
Calculated R^2	0.79
Intraclass correlation	**Percent variation**
Country-level	16%
Municipality-level	40%
Observation-level	44%
Estimated random intercepts for country dummy variables	**Mean (Std Dev)**
El Salvador	0.02
	(0.05)
Guatemala	-0.01
	(0.04)
Honduras	-0.01
	(0.03)

*Significance level: p<0.01***; p<0.05***

Discussion

In this study we assessed and described land changes in an understudied transboundary area – the Trifinio Region – by assessing LCLUC at different boundaries and estimating direct drivers of these changes. Below we discuss the significance and implications of these findings for understanding transboundary forest change.

General trends and drivers of land-cover and land-use change

Annual rates of forest loss in Trifinio measured highest in Guatemala (2%), followed by Honduras (1.4%), and El Salvador (1.2%). At the national scale, the average annual rate of forest loss (nationally) over the period 1986-2010 was 1.3% for Guatemala (FAO, 2010a), 2.2% for Honduras (FAO, 2010c), and 1.4% for El Salvador (FAO, 2010b). Although, there are slightly different methods and periods under study (four more years considered in our study period than the FAO statistics), the rates of change were similar for the transboundary area of El Salvador relative to nationwide data but above average in Guatemala, and below average in Honduras. Honduras, though, had the largest area of forest loss (575 km^2), whereas the area of forest loss was 357 km^2 in Guatemala and only 72 km^2 in El Salvador. This suggests considerable variability in forest loss in the region, and further work to assess whether the transboundary designation has any influence on forest change patterns in Trifinio is warranted.

Overall, forest loss in Trifinio follows similar tropical land-use conversions from forest to economic land-use opportunities, such as coffee and cattle (Table 3). Coffee represents a large part of the export economy in each of these countries. Historically, coffee productivity intensified with technological and transport improvements. From 2000-2010, the coffee export average as a share of total export earnings was 9% for El Salvador, 12% for Guatemala, and 20% for Honduras (ICO, 2014). Nearly 500,000 pickers lost their jobs (Luttinger & Dicum, 2006) in the early 2000s when prices dropped. Cattle ranching and coffee production provide important sources of income to landholders, and changes in these land-uses are linked to market conditions that foster expansion of pasture or coffee plantations, and subsequently, when prices are low, can cause unemployed labourers to expand agricultural production on their own lands to meet subsistence needs. The consummate poverty reduction strategy in pursuit of socioeconomic benefits from coffee and

other crops leads to high percentages of forest conversion with negative environmental consequences, not unlike outcomes in other regions where forest reduction comes as a trade-off for rural development (Soares-Filho et al., 2004; Müller & Munroe, 2005). In Honduras in 1987, road-building subsidies were decreed to incentivize coffee-production, resulting in minor producers expanding production to take advantage of the subsidy (Tucker, 2008). This outcome highlights the inherent trade-off in sustainable development and conservation faced by transboundary conservation areas.

Table 1.3. Indirect drivers of forest change that operate at multi-levels.

Major factors	Examples of specific indirect drivers
Institutional	Land tenure system
	Establishment of protected areas
	Land or agrarian reform (1970s-1990s)
	Transboundary Agreement (Plan Trifinio)
	Transboundary protection (MTPA[1]) since 1987
	Transboundary Reserve of the Biosphere (Trifinio-Fraternidad) since 2011
Socio-political triggers	Wars (between 1970s to 1990s)
	Peace agreements (late 1980s, early 1990s)
Market-oriented	Change in commodity prices
	Coffee prices high years (1977, 1986, 1995, 1997, 2011)
	Coffee crisis in early 1990s and 2000s
	Cattle prices high years (1974, 1980s, 1992, early 2010s)
	Cattle prices low years (1976 and late 1990s)
Technological	Improvements in equipment that affect agricultural intensification, agricultural expansion, and deforestation
	Access to machinery and transportation of produce
Demographic	Changes in population
	Immigration (due economic opportunities)
	Outmigration (due to war or constrains)
Biophysical	Natural events (e.g., hurricanes, fire)
	Crop diseases

There was considerable heterogeneity in observed land changes across countries and sub-nationally. In our regression model of direct drivers of change, we found that accessibility and population were significantly correlated with forest loss, and that there was more heterogeneity in deforestation at the sub-national than national level. Demographic factors such as migration from rural areas to urban centers may have had an effect on land-use changes in Trifinio. Blackman, Ávalos-Sartorio, & Chow (2007) noted that in the 1990s northern El Salvador (corresponding to the Trifinio region) was characterized by civic unrest and poverty after the civil wars. Massive internal migration to the southwest and greater metropolitan areas near the capital of the country also took place during that time (Edelman, 2008). To our knowledge the linkages between migration and land changes in Trifinio have not been studied. However, because forest regrowth may have resulted from decreased pressure on forestlands brought about by constrained agricultural activity, it is reasonable that available agricultural area may have been related to the same process of migration-and-repopulation-after-strife that occurred in the Peten Region (Carr, 2008b). Trifinio is affected by human, landscape, and climate-based environmental challenges as elsewhere in Central America, including insufficient arable land, located on steep, forested slopes, farmed by a burgeoning rural poor (i.e., ecological marginalization) (Artiga, 2003; CATIE, 2003; GEF, 2006; US Agency for International Development [USAID], 2011; Castellanos et al., 2013).

At the sub-national level, the democratization process and development of new forms of sub-national governance developed in the late 1990's and may have affected land outcomes, with an unintended consequence of higher rates of forest change in some places. Some of the most important factors are associated with institutional changes around land reform. Land tenure in Trifinio includes private, state-owned, and unclaimed lands, and common property resources. Agrarian land reforms, linked to environmental governance by institutional changes related to land, are encapsulated in Spanish literature as '*reforma agraria*'; Diskin noted that the reforms "endeavour to reshape tenure systems correcting conditions fostering inequality, political imbalance, and poverty by restructuring ownership and fostering use changes for development" (as cited in Thiesenhusen, 1989, p. 438). Agrarian reforms implemented from 1970-1990 changed land tenure dramatically (Williams, 1986; Bulmer-Thomas, 1987), and are linked to other indirect drivers of LCLUC such as socio-political trigger events; specifically, the civil wars and the peace agreement that

followed. In general, regional documents describe scarce local presence of land governance in Trifinio. Governance and political stability are not sufficient conditions to drive change, however, but if coupled with global or locally higher prices (e.g. coffee, livestock) the effects of rural development policies on deforestation can be exacerbated. Additional investigation of nested governance and sub-national policies within the transboundary governance area are needed to understand fully how these institutional and governance changes relate to changes in land-cover.

Several researchers reported forest-woodland resurgence in El Salvador during 2001-2010, resulting from effects of globalization and integration through the combination of remittances, off-farm income, commodity markets, environmentalism, and agrarian reform (Hecht & Saatchi, 2007; Castaneda, 2009; Hecht, 2010; Herrador Valencia, Boada i Juncà, Varga Linde, & Mendizábal Riera, 2011). In addition, other indirect drivers associated with LCLUC in this region that are likely to have affected Trifinio include conversion from sun to shaded coffee, severe regional pests such as the coffee berry borer and leaf rust, and natural events such as hurricanes and wildfire (Eakin, Tucker, & Castellanos, 2006; Vega, Infante, Castillo, & Jaramillo, 2009; Castellanos et al., 2013; Haggar, Medina, Aguilar, & Munoz, 2013; Schmitt-Harsh, 2013; El Mundo, 2015).

As noted above, increases in forest coverage were found for both El Salvador (8%-9% from 1986-2001), and Honduras (62%-65%, from 2001-2010), and these could be related to increased use of traditional coffee varieties that are shaded by native forest vegetation, and hidden during Landsat analyses (Montenegro & Atwood, 2010; Ortega-Huerta, Komar, Price & Ventura, 2012; Schmitt-Harsh, Sweeney, & Evans, 2013). Very high-resolution (VHR) remote sensing products (e.g. IKONOS, Worldview 2, and aerial orthophotos) can detect shaded coffee formations; unpublished results show individual shade trees, as well as rows of coffee, for this study region (P. Schlesinger, unpublished data). Shaded coffee has been largely promoted because of its benefit to biodiversity conservation and livelihoods (Rice, 2010), though there is argument about its efficacy (Calvo & Blake, 1998; Rappole, King, & Rivera, 2003; Tejeda-Cruz, Silva-Rivera, Barton, & Sutherland, 2010).

Overall, scientists conducting remote sensing and mapping of tree cover in the tropics have made many important strides over the lifetime of the Landsat satellite program (Loveland & Dwyer, 2012; Hansen et al., 2013; Roy et al., 2014), such as those by the

LCLUC global assessment teams (e.g. Hansen et al., 2013; Kim et al., 2014; and, GLCF/GSFC, 2014). These global products were not designed to map human land-use, though they have traditionally captured large changes due to industrial agriculture and forest clearcut activity (Defries et al., 2005). Their format of hard classifications of strictly 30 x 30 meter forest and nonforest pixels consistently underestimate forest losses in tropical Latin America (Kim et al., 2015) by obscuring tropical smallholder cropping and other regionally important economic land-uses (such as pasture and agroforestry) that remain a significant and important challenge to estimate (Lu, Batistella, Moran, & de Miranda, 2005; Jain, Mondal, DeFries, Small, & Galford, 2013). The large difference found between our regional findings of deforestation and the latest global efforts (e.g. Hansen et al., 2013) is likely due to inadequate detection of smallholder land-uses, particularly pastures and coffee agroforests. Major land-use changes in the tropics, from small-scale sun and shaded coffees, cocoas, and other forms of agroforestry, and woody vegetation, termed 'trees outside of forests' (TOF) by FAO (Schnell, 2015), are often only visible at higher spatial resolutions beneath canopies and without cloud cover. Thus, higher resolution (and radar) satellite data may provide further insight to tropical land-use change (Ryan et al., 2012, Vaglio Laurin et al., 2014). Devoting attention to cost-effective and accurate mapping of tree-based agricultural and agroforestry systems would therefore be of great benefit to many LCLUC-related studies in the tropics. Trends and drivers of changes within and around the MTPA.

Forest loss can be seen as a positive outcome of agriculture and livestock expansion, because other benefits and opportunities can be realized with the right policies in place for forest protection (e.g. PAs, biological corridors, and incentives). Finding this balance is fundamental for policy design. Thus, it is important to understand the roles of socioeconomic dynamics and biophysical drivers in determining LCLUC, especially in and around PAs, whether these are domestic or transnational (DeFries, Hansen, Turner, Reid, & Liu, 2007).

The spatial biophysical composition of the landscape of this region (e.g. amount of forest, elevation, and slope) may also define what specific human activities (e.g. protection of land, forest conversion) slow or intensify the process of change from direct drivers. Our analysis of the MTPA core zone finds that forest cover is maintained throughout the study period relative to surrounding buffer zones. As compared to other transboundary protected areas in Central America (e. g. La Amistad), our observations of general trends indicated that

the tri-national governance of this protected area appears to have had positive benefits for forest protection within its boundaries.

The tapering of deforestation as one moves from the core Montecristo park boundary to further away might be attributed to: a) better land for coffee existing at the mid-to-higher altitude lands that are closer to the park boundary (Blackman et al., 2008; Fischer & Victor, 2014); b) the fact that there was little suitable land left for agroforestry coffee that had not already been cultivated further from the park as of 1986; or c) the park itself may have led to 'spillovers' of land-use pressure at mid-to- higher elevation directly around its boundaries (Joppa & Pfaff, 2010). The large annual rate of change observed for the urban class corresponded geographically to the areas just beyond 4 km buffer that saw significant disturbance. The very high-level disturbance immediately outside the park, and the ability to maintain forest cover within the park core over time is notable and suggests that its protected status constrained LCLUC within the park over the study period.

Conducting an assessment of the causal impact of the protected area on deforestation was beyond the scope of this study, and so we cannot conclude that protection in forest cover is 'additional' to what would have occurred without the protected area (Joppa & Pfaff, 2010). Further analysis is needed to elucidate the causal impact of the park on deforestation by more carefully controlling for confounding factors (Joppa & Pfaff 2010).

Conclusion

The creation of the Trifinio Region and its MTPA appears to have safeguarded forest in the vicinity of the tri-national park. However, across the larger region, key drivers of LCLUC during the study period are similar to those found elsewhere in Central America – with agricultural expansion of coffee, pasture, and other crop production – suggesting that there is little detectable difference from the transboundary governance arrangement on LCLUC. Additional research is needed, however, to test the causal relationships between transboundary governance, including the MTPA, within Trifinio and its impact of natural resources management in comparison with efforts outside of the region. While we observed considerable heterogeneity in land changes across countries, including within each country's MTPA core and buffer zones, our regression analysis suggested that more heterogeneity occurred in forest loss patterns at the sub-national level within each of the Trifinio countries

than across countries. At the sub-national level, our regression model of direct drivers of change suggested that factors like road access to forest resources and population were significantly correlated with forest loss. There are several important indirect drivers of land-cover change in Trifinio, such as recent land reforms and rapid changes in commodity prices, linked to these observed direct drivers. Finally, our analysis revealed large differences in deforestation between global scale forest assessments and measured local changes, highlighting the need for further research in mapping tropical tree crops and shaded agroforestry.

Acknowledgements

We dedicate this research to the memory of Juan Carlos Montufar Celada, Technical Manager of the Trifinio Region, whose life's work was devoted to peaceful transboundary conservation and cooperation. We gratefully acknowledge the assistance of GIZ Forests and Water Program, US Fish and Wildlife, Tri-national Commission of the Trifinio Plan, and the Center for Tropical Agricultural Research and Education (CATIE)'s Mesoamerican Agroenvironmental Program (MAP Norway) for their assistance in field access and data collection.

Funding

This work was funded by NASA LCLUC Grant NNX13AC70G (K. Jones, PI) and MICITT/CONICIT in Costa Rica. The authors of this paper have no financial and/or business interests in the Trifinio Region nor any companies that may be affected by the research reported in the enclosed paper.

References

Angelsen, A., & Kaimowitz, D. (1999). Rethinking the causes of deforestation: lessons from economic models. *The World Bank Research Observer, 14*(1), 73–98. Retrieved from http://wbro.oxfordjournals.org/content/14/1/73.shortAnselin, L. (2002). Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, *27*(3), 247–267. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169515002000774

Anselin, L. (2002). Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, *27*(3), 247–267. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169515002000774Artiga, R. (2002). Límites y potencialidades para la gestión integrada de los recursos hídricos de cuencas transfronterizas de Centroamérica. *Estudio de Caso: Gestion de La Cuenca Compartida Del Rio Lempa (Guatemala, El Salvador, Honduras, Plan Trifinio) Y Su replicabilidad–Fundación Del Servicio Exterior Para La Paz Y La Democracia Fundepam. San Salvador, El Salvador*, *68*.

Artiga, R. (2003). *The Case of the Trifinio Plan in the Upper Lempa: Opportunities and Challenges for the Shared Management of Central American Transnational Basins*. Unesco. Retrieved from http://webworld.unesco.org/water/wwap/pccp/cd/pdf/case_studies/the_case_of_the_trifinio_plan_in_the_upper_lempa_2.pdf

Blackman, A., Ávalos-Sartorio, B., & Chow, J. (2007). Shade coffee & tree cover loss: lessons from El Salvador. *Environment: Science and Policy for Sustainable Development*, *49*(7), 22–33. Retrieved from http://www.tandfonline.com/doi/pdf/10.3200/ENVT.49.7.22-33

Blackman, A., Ávalos-Sartorio, B., & Chow, J. (2008). Land cover change in mixed agroforestry: shade coffee in El Salvador. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1272669

Bulmer-Thomas, V. (1987). *The political economy of Central America since 1920*. Cambridge; New York: Cambridge University Press.

Calvo, L., & Blake, J. (1998). Bird diversity and abundance on two different shade coffee plantations in Guatemala. *Bird Conservation International*, *8*(03), 297–308. http://doi.org/10.1017/S0959270900001945

Carius, A. (2007). *Environmental peacebuilding: Environmental cooperation as an instrument of crisis prevention and peacebuilding: conditions for success and constraints*. Berlin: Adelphi Consult.

Carr, D. L. (2008a). Farm households and land use in a core conservation zone of the Maya Biosphere Reserve, Guatemala. *Human Ecology*, *36*(2), 231–248. http://doi.org/10.1007/s10745-007-9154-1

Carr, D. L. (2008b). Migration and tropical deforestation: Why population matters. *Progress in Human Geography*, *33*(2), 355–378.

Castaneda, H. (2009). *Analysis of the spatial dynamics and drivers of forest cover change in the Lempa River Basin of El Salvador*. [Gainesville, Fla.]: University of Florida. Retrieved from http://purl.fcla.edu/fcla/etd/UFE0024235

Castellanos, E. J., Tucker, C., Eakin, H., Morales, H., Barrera, J. F., Diaz, R., & Rethinking Integrated Assessments and Management Projects in the Americas. (2013). Assessing the adaptation strategies of farmers facing multiple stressors: Lessons from the Coffee and Global Changes project in Mesoamerica. *Environmental Science and Policy*, *26*, 19–28.

CATHALAC. (2011). *Cobertura y Uso de la Tierra de la Región del Trifinio: Estudio de los años 1986, 2001 y 2010 mediante métodos de teledetección*. Panamá.

CATIE (Centro Agronómico Tropical de Investigación y Enseñanza). (2003). *Perfil de proyecto de manejo de la subcuenca del Río Copán para la protección del parquet arqueológico de Copán Ruinas*. Honduras.

Celata, F., Coletti, R., & Sanna, V. S. (2013). La cooperación transfronteriza en la región del Trifinio y la difusión de modelos europeos de gobernanza de las fronteras en América Latina. *Si Somos Americanos. Revista de Estudios Transfronterizos*, *13*(2), 165–189. Retrieved from http://www.scielo.cl/pdf/ssa/v13n2/art08.pdf

Choi, C. Q. (2008, September 15). Drug traffickers and other outlaws endanger forest preservation efforts: Illegal ranching and illicit activities hamper forest conservation

efforts. *Scientific American*. Retrieved from http://www.scientificamerican.com/article/drug-traffickers-endanger-preservation/

Center for International Earth Science Information Network - CIESIN - Columbia University, & Information Technology Outreach Services - ITOS - University of Georgia. (2013). Global Roads Open Access Data Set, Version 1 (gROADSv1). NASA Socioeconomic Data and Applications Center (SEDAC). Retrieved from http://dx.doi.org/10.7927/H4VD6WCT. Accessed March 15, 2015.

CTPT. (1997). *Tratado entre las Repúblicas de El Salvador, Guatemala y Honduras para la ejecución del Plan Trifinio*. El Salvador, Guatemala and Honduras.

CTPT. (2011). *Estado de la Región Trifinio 2010. Datos socioeconómicos y ambientales de los municipios.* GIZ y CTPT. November 2011.

DeFries, R., Asner, G., Achard, F., Justice, C., Laporte, N., Price, K., … Townshend, J. (2005). Monitoring tropical deforestation for emerging carbon markets. *Tropical Deforestation and Climate Change*, 35–44. Retrieved from http://cmapspublic3.ihmc.us/rid=1H42JRS7V-275MYTJ-F3N/4930_TropicalDeforestation_and_ClimateChange.pdf#page=35

DeFries, R., Hansen, A., Turner, B. L., Reid, R., & Liu, J. (2007). Land Use Change Around Protected Areas: Management to Balance Human Needs and Ecological Function. *Ecological Applications*, *17*(4), 1031–1038. http://doi.org/10.1890/05-1111

Dinerstein, E. (Ed.). (1995). *A conservation assessment of the terrestrial ecoregions of Latin America and the Caribbean*. Washington, D.C: World Bank.

Eakin, H., Tucker, C., & Castellanos, E. (2006). Responding to the coffee crisis: a pilot study of farmers' adaptations in Mexico, Guatemala and Honduras. *Geographical Journal*, *172*(2), 156–171. http://doi.org/10.1111/j.1475-4959.2006.00195.x

Eastman, J. R. (1989). IDRISI: A geographic information system for international development. Presented at the Conference on Information Technologies for Developing Countries, California: University of Southern California.

Edelman, M. (2008). Transnational organizing in agrarian Central America: histories, challenges, prospects. *Journal of Agrarian Change*, *8*(2-3), 229–257. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1471-0366.2008.00169.x/pdf

El Mundo. (2015, June 23). Expertos analizan impacto de la roya en Guatemala | Diario El Mundo [Online News El Salvador]. Retrieved November 16, 2015, from http://elmundo.sv/expertos-analizan-impacto-de-la-roya-en-guatemala/

Environmental Systems Research Institute (ESRI). (2012). *ArcGIS Release 10.1*. Redlands, CA.

FAO. (1995). *Global forest land-use change 1990-2005* (No. FAO Forestry Paper 124). Rome: Food and Agriculture Organization of the United Nations FAO. Retrieved from http://www.fao.org/docrep/007/v5695e/v5695e00.htm

FAO. (2010a). Evaluación De Los Recursos Forestales Mundiales 2010. Informe Nacional. El Salvador. FRA 061. Roma. Seccion 1.4. Retrieved November 16, 2015, from http://www.fao.org/docrep/013/al520S/al520S.pdf

FAO. (2010b). Evaluación De Los Recursos Forestales Mundiales 2010. Informe Nacional. Guatemala. FRA 084. Roma. Seccion 1.4. Retrieved November 16, 2015, from http://www.fao.org/docrep/013/al520S/al520S.pdf

FAO. (2010c). Evaluación De Los Recursos Forestales Mundiales 2010. Informe Nacional. Honduras. FRA 091. Roma. Seccion 1.4. Retrieved November 16, 2015, from http://www.fao.org/docrep/013/al527S/al527S.pdf

FAO. (2010d). *Global Forest Resources Assessment 2010. Main Report* (FAO Forestry Paper No. 163). Rome: Food and Agriculture Organization of the United Nations FAO. Retrieved from http://www.fao.org/docrep/013/i1757e/i1757e.pdf

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., … others. (2007). The shuttle radar topography mission. Reviews of Geophysics, 45(2). Retrieved from http://onlinelibrary.wiley.com/doi/10.1029/2005RG000183/full

Fischer, E. F., & Victor, B. (2014). High-End Coffee and Smallholding Growers in Guatemala. *Latin American Research Review*, *49*(1), 155–177. Retrieved from http://muse.jhu.edu/journals/latin_american_research_review/v049/49.1.fischer.html

GEF. (2006). *Project document of the integrated management of the Montecristo trinational protected area*.

Geist, H. J., & Lambin, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *BioScience, 52*(2), 143. Retrieved from

http://ida.lib.uidaho.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=voh&AN=6057112&site=ehost-live&scope=site

German Society for International Cooperation (GIZ). (2011). *Estado de la región trifinio 2010: Datos socioeconómicos y ambientales de los municipios*. German Society for International Cooperation (GIZ).

Girot, P. O. (1997). Border Regions, Integration, and Transborder Conservation Initiatives in Central America. In P. Ganster (Ed.), *Borders and Border Regions in Europe and North America*. San Diego, California,USA: San Diego State University Press. Retrieved from http://books.google.com/books?id=NJl1d7r728IC

Global Land Cover Facility (GLCF), & Goddard Space Flight Center (GSFC). (2014). *GLCF Forest Cover Change 2000 2005, Global Land Cover Facility*. Maryland: University of Maryland, College Park.

Graham, D. (2002). *Mesoamerican Biological Corridor: Mexico to Panama* (Case Study) (p. 2). IUCN. Retrieved from http://www.tbpa.net/docs/62_Meso_American_Biological_Corridor.pdf

Greenbaum, E., & Komar, O. (2005). Threat Assessment and Conservation Prioritization of the Herpetofauna of El Salvador. *Biodiversity and Conservation*, *14*(10), 2377–2395. http://doi.org/10.1007/s10531-004-1670-3

Haggar, J., Medina, B., Aguilar, R. M., & Munoz, C. (2013). Land use change on coffee farms in southern Guatemala and its environmental consequences. *Environmental Management*, *51*(4), 811–823. http://doi.org/10.1007/s00267-013-0019-7

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., … Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853. http://doi.org/10.1126/science.1244693

Hecht, S. (2010). The new rurality: Globalization, peasants and the paradoxes of landscapes. *Land Use Policy*, *27*(2), 161–169. http://doi.org/10.1016/j.landusepol.2009.08.010

Hecht, S. B., & Saatchi, S. S. (2007). Globalization and forest resurgence: Changes in forest cover in El Salvador. *Bioscience*, *57*(8), 663–672. Retrieved from http://bioscience.oxfordjournals.org/content/57/8/663.short

Herrador Valencia, D., Boada i Juncà, M., Varga Linde, D., & Mendizábal Riera, E. (2011). Tropical forest recovery and socio-economic change in El Salvador: An opportunity

for the introduction of new approaches to biodiversity protection. *Applied Geography*, *31*(1), 259–268. http://doi.org/10.1016/j.apgeog.2010.05.012

International Coffee Organization. (2014). World Coffee Trade. Retrieved from http://www.ico.org/trade_e.asp?section=About_Coffee

InterAmerican Development Bank. (2010). Coming together in the Trifinio region of El Salvador, Guatemala and Honduras. Retrieved from http://www.impactalliance.org/file_download.php?location=S_U&filename=12684422775IDB_-_Coming_together_in_the_Trifinio_region_of_El_Salvador%2C_Guatemala_and_Honduras.pdf

IPS. (2002, August 2). Meso-America Being Left Treeless | Inter Press Service. Retrieved November 16, 2015, from http://www.ipsnews.net/2002/08/meso-america-being-left-treeless/

Jain, M., Mondal, P., DeFries, R. S., Small, C., & Galford, G. L. (2013). Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sensing of Environment*, *134*, 210–223. http://doi.org/10.1016/j.rse.2013.02.029

Joppa, L. N., Loarie, S. R., & Pimm, S. L. (2008). On the protection of "protected areas." *Proceedings of the National Academy of Sciences*, *105*(18), 6673–6678. http://doi.org/10.1073/pnas.0802471105

Joppa, L., & Pfaff, A. (2010). Reassessing the forest impacts of protection. *Annals of the New York Academy of Sciences*, *1185*(1), 135–149. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2009.05162.x/fullKim, D.-H., Sexton, J. O., Noojipady, P., Huang, C., Anand, A., Channan, S., … Townshend, J. R. (2014). Global, Landsat-based forest-cover change from 1990 to 2000. *Remote Sensing of Environment*. http://doi.org/10.1016/j.rse.2014.08.017

Kim, D.-H., Sexton, J. O., Noojipady, P., Huang, C., Anand, A., Channan, S., … Townshend, J. R. (2014). Global, Landsat-based forest-cover change from 1990 to 2000. *Remote Sensing of Environment*. http://doi.org/10.1016/j.rse.2014.08.017

Kim, D.-H., Sexton, J. O., & Townshend, J. R. (2015). Accelerated deforestation in the humid tropics from the 1990s to the 2000s: ACCELERATED PAN-TROPICAL

DEFORESTATION. *Geophysical Research Letters*, *42*(9), 3495–3501. http://doi.org/10.1002/2014GL062777Komar, O. (2002). Birds of Montecristo National Park, El Salvador. *Ornitol. Neotrop*, *13*, 167–193. Retrieved from http://132.248.13.1/pdf/links/neo/rev13/vol_13_2/orni_13_2_167-194.pdf

Komar, O. (2002). Birds of Montecristo National Park, El Salvador. *Ornitol. Neotrop*, *13*, 167–193. Retrieved from http://132.248.13.1/pdf/links/neo/rev13/vol_13_2/orni_13_2_167-194.pdf

Lambin, E. F., & Geist, H. J. (2006). *Land-use and land-cover change: local processes and global impacts*. New York: Springer.

Lambin, E. F., Meyfroidt, P., Rueda, X., Blackman, A., Börner, J., Cerutti, P. O., … Lister, J. (2014). Effectiveness and synergies of policy instruments for land use governance in tropical regions. *Global Environmental Change*, *28*, 129–140.

Lawrence D., & Vandecar K. (2015). Effects of tropical deforestation on climate and agriculture. *Nat. Clim. Change Nature Climate Change*, *5*(1), 27–36. Retrieved from http://www.nature.com/nclimate/journal/v5/n1/full/nclimate2430.html

Loveland, T. R., & Dwyer, J. L. (2012). Landsat: Building a strong future. *Remote Sensing of Environment*, *122*, 22–29. http://doi.org/10.1016/j.rse.2011.09.022

Lu, D., Batistella, M., Moran, E. F., & de Miranda, E. E. (2005). A comparative study of Terra ASTER, Landsat TM, and SPOT HRG data for land cover classification in the Brazilian Amazon. In *The 9th World Multi-Conference on Systematics, Cybernetics, and Informatics (WMSCI 2005). Orlando, FL: International Institute of Informatics and Systematics* (pp. 411–416). Retrieved from https://www.msu.edu/~moranef/documents/05-11_AComparativeStudy.pdf

Luttinger, N., & Dicum, G. (2006). *The coffee book: anatomy of an industry from crop to the last drop*. New York: New Press.

Mendoza, E., Fuller, T. L., Thomassen, H. A., Buermann, W., RamíRez-MejíA, D., & Smith, T. B. (2013). A preliminary assessment of the effectiveness of the Mesoamerican Biological Corridor for protecting potential Baird's tapir (*Tapirus bairdii*) habitat in southern Mexico. *Integrative Zoology*, *8*(1), 35–47. http://doi.org/10.1111/1749-4877.12005

Meyfroidt, P. (2015). Approaches and terminology for causal analysis in land systems science. *Journal of Land Use Science*, 1–27. Retrieved from: http://www.tandfonline.com/doi/abs/10.1080/1747423X.2015.1117530

Ministerio de Medio Ambiente y Recursos Naturales de El Salvador (MARN). (2010). *Reserva de la Biósfera Trifinio Fraternidad*. Ministerio de Medio Ambiente y Recursos Naturales de El Salvador. Miranda, J. A., Slowing, K., & Raudales, J. C. (2010). South-south learning in the Trifinio region: Transforming borderlands into areas of peace and development. *Development Outreach*, *12*(2), 30–31.

Montenegro, R., & Atwood, D. (2010). *Identification of Shade Coffee Using Optical / SAR Data Fusion*. Poster presented at the 2010 AGU Meeting of the Americas, Foz do Iguacu, Brazil. Retrieved from http://www.agu.org/meetings/ja10/pdf/AGU_JA10_Program.pdf

Müller, D., & Munroe, D. K. (2005). Tradeoffs between Rural Development Policies and Forest Protection: Spatially Explicit Modeling in the Central Highlands of Vietnam. *Land Economics*, *81*(3), 412–425. http://doi.org/10.3368/le.81.3.412

Nepstad, D., Schwartzman, S., Bamberger, B., Santilli, M., Ray, D., Schlesinger, P., … Rolla, A. (2006). Inhibition of Amazon Deforestation and Fire by Parks and Indigenous Lands. *Conservation Biology*, *20*(1), 65–73. http://doi.org/10.1111/j.1523-1739.2006.00351.x

Oestreicher, J. S., Benessaiah, K., Ruiz-Jaen, M. C., Sloan, S., Turner, K., Pelletier, J., … Potvin, C. (2009). Avoiding deforestation in Panamanian protected areas: An analysis of protection effectiveness and implications for reducing emissions from deforestation and forest degradation. *Global Environmental Change*, *19*(2), 279–291. http://doi.org/10.1016/j.gloenvcha.2009.01.003

OAS. (1993). *Plan Trifinio El Salvador - Guatemala - Honduras 1992*. Washington, D.C., USA: Secretariat General of the Organization of American States (OAS).

Ortega-Huerta, M. A., Komar, O., Price, K. P., & Ventura, H. J. (2012). Mapping coffee plantations with Landsat imagery: an example from El Salvador. *International Journal of Remote Sensing*, *33*(1), 220–242. http://doi.org/10.1080/01431161.2011.591442

Pan, W. K. Y., & Bilsborrow, R. E. (2005). The use of a multilevel statistical model to analyze factors influencing land use: a study of the Ecuadorian Amazon. *Global and Planetary Change*, *47*(2-4), 232–252. http://doi.org/10.1016/j.gloplacha.2004.10.014

Puyravaud, J.-P. (2003). Standardizing the calculation of the annual rate of deforestation. *Forest Ecology and Management*, *177*(1–3), 593–596. http://doi.org/10.1016/S0378-1127(02)00335-3

Rabe-Hesketh, S. & Skrondal, A. (2008). *Multilevel and longitudinal modeling using stata* (2nd ed). College Station, Tex: Stata Press Publication.

Rappole, J. H., King, D. I., & Rivera, J. H. V. (2003). Coffee and conservation III: reply to Philpott and Dietsch. *Conservation Biology*, 1847–1849. Retrieved from http://www.jstor.org/stable/3588931

Rice, R. (2010). The ecological benefits of shade-grown coffee: the case for going bird friendly. *Smithsonian Migratory Bird Center, Washington*. Retrieved from http://nationalzoo.si.edu/scbi/migratorybirds/coffee/bird_friendly/Eco-Report.pdf

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., … Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172. http://doi.org/10.1016/j.rse.2014.02.001

Ryan, C. M., Hill, T., Woollen, E., Ghee, C., Mitchard, E., Cassells, G., … Williams, M. (2012). Quantifying small-scale deforestation and forest degradation in African woodlands using radar imagery. *Global Change Biology*, *18*(1), 243–257. http://doi.org/10.1111/j.1365-2486.2011.02551.x

Sandwith, T. (Ed.). (2001). *Transboundary protected areas for peace and co-operation: based on the proceedings of workshops held in Bormio (1998) and Gland (2000).* Gland, Switzerland: IUCN--the World Conservation Union.

Schmitt-Harsh, M. (2013). Landscape change in Guatemala: Driving forces of forest and coffee agroforest expansion and contraction from 1990 to 2010. *Applied Geography*, *40*, 40–50. http://doi.org/10.1016/j.apgeog.2013.01.007

Schmitt-Harsh, M., Sweeney, S. P., & Evans, T. P. (2013). Classification of coffee-forest landscapes using Landsat TM imagery and spectral mixture analysis. *Photogrammetric Engineering & Remote Sensing*, *79*(5), 457–468. Retrieved from

http://www.ingentaconnect.com/content/asprs/pers/2013/00000079/00000005/art00004

Schnell, S. (2015). Integrating trees outside forests into national forest inventories. Retrieved from http://pub.epsilon.slu.se/12011/

Soares-Filho, B., Alencar, A., Nepstad, D., Cerqueira, G., Vera Diaz, M. del C., Rivero, S., … Voll, E. (2004). Simulating the response of land-cover changes to road paving and governance along a major Amazon highway: the Santarem-Cuiaba corridor. *Global Change Biology*, *10*(5), 745–764. http://doi.org/10.1111/j.1529-8817.2003.00769.x

Stevens, C. (2013). Resilient Environmental Governance: Protecting Changing Ecosystems Through Multilevel Governance. Retrieved from http://scholarworks.umass.edu/open_access_dissertations/842/

Tejeda-Cruz, C., Silva-Rivera, E., Barton, J. R., & Sutherland, W. J. (2010). Why shade coffee does not guarantee biodiversity conservation. *Ecology and Society*, *15*(1), 13. Retrieved from http://www.ecologyandsociety.org/vol15/iss1/art13/

Thiesenhusen, W. C. (1989). *Searching for agrarian reform in Latin America.* Boston: Unwin Hyman. Retrieved from http://pdf.usaid.gov/pdf_docs/PNABC444.pdf

Tucker, C. M. (2008). Coffee production and communal forests in Honduras: Adaptation and resilience in a context of change. Presented at the IASC 2008, Cheltenham, England: International Association for the Study of the Commons. Retrieved from http://iasc2008.glos.ac.uk/conference%20papers/papers/T/Tucker_228601.pdf

Turner, B. L., Lambin, E. F., & Reenberg, A. (2007). The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences*, *104*(52), 20666–20671. http://doi.org/10.1073/pnas.0704119104

United Nations Environment Program & Permanent Commission for the South Pacific. (2006). *Eastern Equatorial Pacific, GIWA Regional assessment 65*. Sweden: UNEP & CPPS.

United States Agency International Development (USAID). (2011). *USAID country profile: Property rights and resource governance: El Salvador*. United States Agency International Development (USAID). Retrieved from http://usaidlandtenure.net/sites/default/files/country-profiles/full-reports/USAID_Land_Tenure_El_Salvador_Profile_0.pdf

Vance, C., & Iovanna, R. (2006). Analyzing spatial hierarchies in remotely sensed data: Insights from a multilevel model of tropical deforestation. *Land Use Policy*, *23*(3), 226–236.

Vega, F. E., Infante, F., Castillo, A., & Jaramillo, J. (2009). The coffee berry borer, Hypothenemus hampei (Ferrari)(Coleoptera: Curculionidae): a short review, with recent findings and future research directions. *Terrestrial Arthropod Reviews*, *2*(2), 129–147. Retrieved from http://www.ncaur.usda.gov/SP2UserFiles/person/5818/TerrestrialArthropodReviews.pdf

Williams, R. G. (1986). *Export agriculture and the crisis in Central America*. Chapel Hill: University of North Carolina Press.

Zbicz, D. C. (1999). *Transboundary cooperation in conservation: A global survey of factors influencing cooperation between internationally adjoining protected areas* (Ph.D.). Duke University, Ann Arbor. Retrieved from ProQuest Dissertations & Theses Full Text; ProQuest Dissertations & Theses Global. (304503067)

Zbicz, D. C. (2001). Crossing international boundaries in park management—a survey of transboundary cooperation. In *Crossing Boundaries in Park Management: Proceedings of the 11th Conference on Research and Resource Management in Parks and on Public Land*. Hancock, Michigan: The George Wright Society. Retrieved from http://www.georgewright.org/33zbicz.pdf

CHAPTER 2: INVESTIGATING FOREST COVER TRANSITIONS OVER 30 YEARS WITH LANDSAT IMAGES IN THE TRIFINIO REGION

Abstract

International development to reduce poverty at the same time as conserving natural resources requires a good understanding of vegetative transitions over the recent past, how and when they increase, decrease, and when they remain constant. A database of percent greenness derived from normalized difference vegetation index data was constructed from 30 years of medium resolution-satellite imagery to understand land cover and land use changes (across countries, counties, and PAs) of a unique Central American transboundary region. A novel method for evaluating temporal trends was applied using short time series analysis tools; temporal trends were largely found to be flat. Proximity to PAs showed a relation to greenness inside the Region. Despite centralized environmental conservation and management, vegetative growth varies widely across the region. Regional regrowth or resurgence of natural forest cover was not observed.

Keywords: Coffee, deforestation, land use, NDVI, protected area, transboundary, Trifinio

Resumen

El desarrollo internacional para reducir la pobreza al mismo tiempo que la conservación de los recursos naturales requiere una buena comprensión de las transiciones vegetativas en el pasado reciente, cómo y cuándo aumentan, disminuyen y cuándo permanecen constantes. Una base de datos de porcentaje de verdor derivado de datos de índice de vegetación de diferencia normalizada se construyó a partir de 30 años de imágenes satelitales de resolución media para comprender la cobertura terrestre y los cambios en el uso de la tierra (en países, condados y áreas protegidas) de una región transfronteriza única de Centroamérica. Un metodo novedoso para la evaluacion de la tendencia temporal fue aplicado usando herramientas de análisis de series temporales cortas; las tendencias temporales fueron descubiertas en gran parte que eran planas. La proximidad a las áreas protegidas mostró una relación con el verdor dentro de la Región. A pesar de la gestión y conservación ambiental

centralizada, el crecimiento vegetativo varía ampliamente en la región. No se observó rebrote regional o resurgimiento de la cubierta forestal natural.

Palabras claves: Café, deforestación, uso del suelo, área protegida, NDVI, índice de vegetación, verdor, transfronterizo, Trifinio

Introduction

Improved understanding of the complexity of the dynamics of vegetated land cover and land use changes (LCLUCs) for the monitoring of transboundary PAs (PAs) supports a wide range of decisions. The dynamics of LCLUC patterns are extremely important for managers and planners of natural resource conservation of nationally adjoining parks, and wildlife refuges that span national borders across the Americas. The notion of transboundary and adjoining parks offers direct potential for reduction of international conflict, and increased stability, peace and conservation, because nations need to work together to collectively maintain forest and cropland productivity (Schlesinger et al, 2017). Surrounding Central America's Trifinio Region, a 7400-km^2 transboundary region at the join of the boundaries of Guatemala, Honduras, and El Salvador, human land use has been increasing levels of landscape fragmentation largely resulting in forest fragments existing in a matrix of agricultural lands.

Significant time and financial resources have been expended since the late 1980s in the continued development of the Trifinio Plan. However, assessment of whether the Plan has been successful from the perspective of conservation, LCLUC, international development, poverty reduction, and water source protection has been limited (Artiga, 2003). The region came together in the aftermath of civil wars, in common national interest to protect the water supply source of three important watersheds, the Ulua, Lempa, and Motagua that supply water to millions of downstream users. Investments in the region are supported by the international community, and its centralized administration is a combined effort of the offices of the Vice Presidents of the three participating countries. It would therefore be expected that the common centralized goals, planning, and conservation management implementation of the participating countries, counties, and PAs would manifest in an increase of greenness across the landscape over time.

Remotely-sensed greenness imaging based on normalized differencing of red and near-infrared optical satellite image bands, such as the Normalized Difference Vegetation Index (NDVI) (Rouse et al, 1974, Deering et al, 1975) enables long term, regional greenness monitoring. Mapping with NDVI allows for LCLUC and degradation to be studied in a consistent manner using multiple satellite sensors (Yengoh et al., 2014). Arithmetic manipulation of the image bands responsible for the ratio of absorption of red light by green chlorophyll to the reflection of near infrared (NIR) light by expanding leaf mesophyll cells highlights differences in vegetative responses to changes in its environment. Early study of the relationships between prairie grass biomass samples and *in situ* spectrometer values (Tucker, 1978) offered an assessment of the various equations developed up to that that time (some of which had been applied to Landsat satellite digital numbers) and concluded that linear combinations of red and infrared radiation were suitable to assess plant canopy photosynthesis. Mathematical equations and relationships investigated following this same logic have led to the creation of nearly two hundred vegetation indices and applications discriminating natural and anthropogenic differences and disturbances (Heinrich et al, 2012). Thus greenness-based land cover and land use transitions can be sampled and used over time in satellite imagery to investigate and measure changes in vegetative cover commensurate with various land uses relating to agriculture, agroforestry, and pasture.

The creation of databases to conduct vegetation analyses in areas once covered by seasonally dry tropical forests with Landsat and other optical remote sensing data can be challenging. Seasonally dry tropical regions are fraught with high levels of cloud cover and cloud shadows, fires, and dynamic water bodies. Recent scholarship, however, has created means of identifying and creating ways to consider some of these dynamic changes in Landsat bands, while detailing the location of cloud cover, shadows, and water (Zhu & Woodcock, 2012). These features can then be removed, and multiple dates of these images can be compared and fused to create more useful data sets for understanding LCLUC over time. Calibration of satellite images is required to prepare vegetation indices for characterization of land uses and their measurement over time, facilitating the search for forest regrowth.

Reforestation has recently become a goal of each of the Trifinio Region's participating countries, however, there has not been a regional study of forest regrowth trends (Figure 2.1).

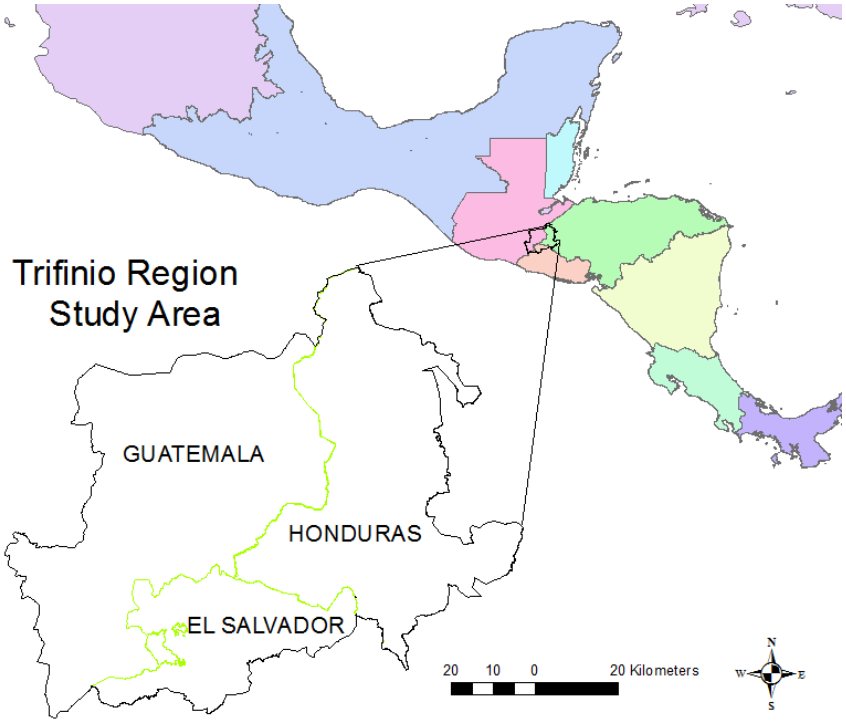


Figure 2.1. The Trifinio study region participating countries are El Salvador, Guatemala, and Honduras.

In Honduras, the government has recently amended mandates and budgeted funds to support reforestation, but the country also directly supports large monocultures (African Palm, coffee, sugar cane, and pasture to support cattle) (USAID, 2014). In 2016, in time for the International Arbor Day, the Honduran government began its new "Honduras plants a life" campaign.

From 2015 to 2017 Guatemala's National Institute of Forest annually began the reforestation campaigns with slogans of "Sowing Traces", "It's time to reforest Guatemala", and "We Will Reforest Guatemala", and goals of replanting 1 million tree seedlings. Contradictions in development policies are rampant, and forestry is no different than other areas. Forestry development activities were created to support pine species, at the same time as rural dwellers were encouraged to use pine species as household fuel. Likewise, in Guatemala, the Constitution's Article 126 (OAS, 1993) declared reforestation and forest conservation matters of national urgency, and there have been efforts to bring these to the forefront with national LCLUC monitoring to accommodate reporting responsibilities for the 2015 Paris Agreement. However, economic development policy has also permitted growth of

the highest national deforestation rates in supporting expansion of African palm, sugar cane, cattle, and narco-farms (Foucart, 2011).

El Salvador's national environmental act was enacted in 1998, but progress in the land use and forest sector has been slow to materialize. The country is participating in UNFCCC preparations to implement the REDD+ mitigation mechanism, but real progress on this front is only expected to materialize this year (CATIE, 2017). The Ministry of Environment and Natural Resources (MARN) is charged with forest and forestry issues in El Salvador; but there is not a tremendous amount of marketable forest cover, as sugar cane and coffee production are significant agricultural players (estimates from unofficial national mapping of land cover and land use (Jimenez, 2014, updated in 2017) show 56% agriculture and 32% forest cover). In April 2016, on Earth Day, MARN launched a reforestation and restoration program in five priority sites, three of which include PAs, wetlands restoration, as well as the promotion of coffee and cocoa (MARN, 2016); a total of 14,000 ha are designated for Trifinio and regions nearby as a part of the Bonn Challenge initiative. Cocoa is preferred over coffee for new agricultural development due to rampant destruction by the coffee rust ('roya' in Spanish). Coffee cooperative growers from El Salvador at an event at CUNORI in Chiquimula reported to us that their government is urging a conversion to cocoa, but has offered no incentives nor supports for the period in which it will take to convert their crops, and many are considering selling their lands.

Forest regrowth and resurgence, and natural reforestation have been noted by published Central American studies, including some local level studies in the Trifinio Region (Redo et al, 2012; Vaca et al, 2012; Bray, 2009; Chowdhury, 2010; Hecht, 2010; Castaneda, 2009; Hecht and Saatchi, 2007; Tucker et al, 2005; Sader et al, 2003; Aguilar, 2002; Southworth and Tucker, 2001, Sader et al, 2001; Turner, 2010). Redo et al (2012), who examined 2000-2010 using Moderate Resolution Imaging Spectrometer (MODIS) for Guatemala and the Trifinio Region showed a marked recovery of conifers, but almost no broadleaf recovery; the study noted Guatemala as dominated by net deforestation. Castaneda (2009) and Hecht and Saatchi (2007) also documented a forest resurgence in El Salvador, but their study areas did not include the Trifinio Region.

Looming increases in satellite image breadth and frequency offer increased ability to study forest cover changes, but some areas still face more considerable challenges in terms of

the number of archived images available for analyses due to incessant cloud cover, limited interest and ability of some nations to map their territories, and higher economic and human priorities. The Trifinio Region has a long rainy season, and poor rural farmers are trying to convert the remaining forest cover to pasture and agroforestry. The contribution of this research is a spatiotemporal assessment of well-known vegetation indices using a novel method to assess trends of forest transition and regrowth inside and outside of a transboundary region, its participating countries, counties, and PAs, that are expected to exhibit common behavior due to common forest governance and development. Specifically, we ask, do trends of greenness transitions vary significantly over the last 30 years when measured across the different countries, counties, and parks that make up the Trifinio Region? Are these trends greater near PAs? Is there a perceived greenness resurgence as reported by regional researchers? And finally, if there is a green-up signal, is it related to abandoned agricultural fields, pastures, or coffee shading?

Methods

We created a time series of remotely-sensed Landsat images to assess human impacted land cover transitions inside and outside of the Trifinio Region. Common activities in time series assessment are to try to sufficiently separate cyclical changes relating to vegetation phenology or cloud cover from true departures from the norm, such as a fire or deforestation event. In this study, however, limitations due to cloud cover necessitated that we examine the earliest portion of the dry season, when vegetation is senescing. Using cloud-free Landsat images between 1986 and 2016, we constructed seven epochs of observations to enable analyses of land cover change (Table 2.1).

Table 2.1. Landsat data selected for use.

LCLUC Epoch	Landsat Satellite	Acquisition Date	Percent Cloud Cover
1986	5	February 2, 1986	20
1986	5	April 7, 1986	0
1986	5	March 25, 1987	10
1991	5	March 4, 1991	2
1991	5	March 20, 1991	0
1996	5	January 29, 1996	43
1996	5	March 17, 1996	6
2001	5	February 11, 2001	13
2001	5	March 15, 2001	6
2001	5	March 31, 2001	0
2003	7	April 14, 2003	1
2011	5	March 27, 2011	5
2011	5	April 28, 2011	2
2016	8	April 9, 2016	5

In preparatory investigation for this study, we used CLASlite software (Asner et al, 2009) with Landsat 5, 7, and 8 images to build a traditional time series of mapped forest polygons with which we could measure and use to understand LCLUC (Schlesinger et al, 2017). CLASlite employs a proprietary form of linear mixed pixel analysis to create fractional forest coverages which has been shown to be very successful for landscape analyses (Kanniah et al, 2016; Tarigan, 2016; Allnutt et al, 2013). However, in the Trifinio Region and its participating countries outside of the region, the native forest cover is widely used to shade coffee production. It was therefore not possible to separate native forest from coffee plantations, nor from forest used for coffee shading using Landsat data. We therefore used NDVI, a simple greenness measure, as an indicator of general vegetation cover change over time.

Data

Data Acquisition

Landsat data of 30m spatial resolution were acquired from United States Geological Survey (USGS)/Earth Resources and Observation Science Center (EROS) for LCLUC analyses. Landsat-derived surface reflectance values were used to develop least-cloud approximately five-year "epochs" of vegetation index for use as a proxy for forest transitions to better understand forest disturbance and regrowth (Table 2.1). These images had already been orthocorrected to the USGS Level 1 Terrain Corrected (L1T) level. In addition, they were radiometrically and atmospherically corrected via the Landsat Ecosystem Disturbance Adaptive Processing System (Masek et al, 2013) and processed to include cloud-shadow mask layers through the USGS EROS Science Processing Architecture Climate Data Record program (Zhu & Woodcock, 2012; USGS, 2014). These data were classified as "provisional" by USGS at the time of download because the processing code was subject to change; all data used in this study were consistently preprocessed using the same date of code (mid-July 2016).

Satellite data for the study were sampled to coincide with the end of the deciduous leaf-off dry season (month of March), to enable acquiring images with limited cloud cover (Table 2.1). Landsat 7 incurred an anomalous scanline corrector failure in May 2003 causing systematic data gaps in ensuing images. Though Year 2006 Landsat 7 images were initially explored for use in the time series, this was not possible and therefore, a cloud-free Landsat 7 scene that had been acquired just prior to the sensor anomaly in 2003 was substituted for the 2006 epoch.

Data Preprocessing

Data from Landsats 4, 7, and 8 comprised five-year epochs (centered on years 1986, 1991, 1996, 2001, 2003/2006, 2011, and 2016) of surface reflectance. Administrative GIS data layers were reprojected from local projections into a common Universal Transverse Mercator (UTM) projection for comparison with Landsat data. Global datasets (i.e. Shuttle Radar Topography Mission (SRTM), Global Administrative Areas (GADM), and World Database on Protected Areas (WDPA)) were incorporated to facilitate characterization of trends among administrative boundaries of nations, counties, and protected areas (PA). Spatial

PA boundary data were gathered during separate field visits (Munoz, 2017) from the individual PA managers who provided and permitted the use of these GIS layers. The Trifinio Region's PAs practice both land and water conservation; however, only PAs concerned with land conservation were selected for use. Spatial databases of agriculture and agroforestry users were downloaded from the region's Trinational Territorial Information System (SINTET geoserver, http://sintet.net) in July 2017. Additional remote sensing data included a set of 113 very-high-resolution (VHR) 2013-2014 Worldview 2 (WV2) satellite images of panchromatic and multispectral data (of 0.5m and 2m resolution respectively), that were used in lieu of orthophotos repeatedly to verify LCLUC activity (e.g. Was a pasture in use or abandoned? Were trees located on the edge of a field or in the middle of a field? and Could suggested new vegetation be seen in the area of a pixels of with increasing slope transitions?).

All remote sensing data underwent pre-processes differing by sensor type and epoch because each epoch was made up of different quantities of inputs, though only a single sensor comprised any one epoch (Figure 2.2). Cloud and shadow polygons from the C Function of Mask (CFMask) layers (Foga et al, 2017; Zhu & Woodcock, 2012) were used to mask out the cloud- and shadow-compromised regions from all processing and analyses streams. Water pixels vary spatially and temporally across the dataset (depending on variable rainfall and subsequent river and lake heights); thus, all water pixels were reclassed to Boolean values and logically summed to create a single water mask. Removing all water pixels over the 30-year period prevented potential water-related false transitions from impacting our analyses.

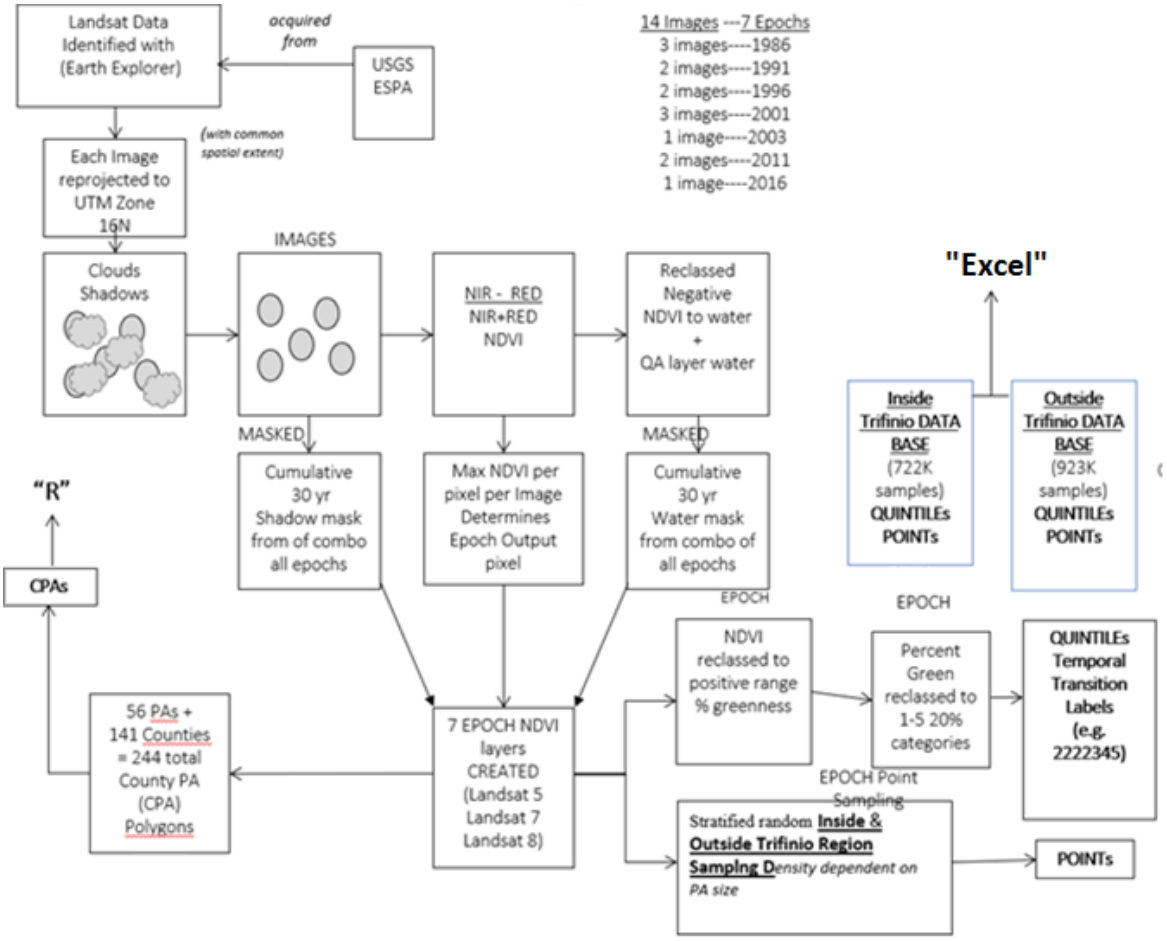


Figure 2.2. Steps followed to produce CPA and point-sampled greenness slope transition classes to assess LCLUC trends.

Preparing the Greenness Database

The original surface reflectance data values range from -10000 to +20000 (USGS, 2017), however, valid data are only held in those values between -10,000 and +10,000. Input integers were rescaled to decimal (real value) reflectance, by multiplying input values by a scale factor of 0.0001. NDVI layers were calculated using the standard formula NIR-RED/NIR+RED (bands 4 and 3 respectively for sensors onboard Landsat 5 and 7, whereas bands 5 and 4 are used with Landsat 8). These data were reclassed to clip the negative range of NDVI values to 0, effectively scaling NDVI to only a positive range. The only reason to keep negative values would have been to capture water, bare soil, snow, ice, cloud, and shadow, but since these values were effectively eliminated with using the CFMask layers and the study is investigating green vegetation, it made no sense to maintain the entire range of

values. From this point on the positive values will be referred to as Percent Greenness (or % greenness).

NDVI layers for each epoch were prepared in five-ranked greenness categories (with linear quintiles of positive NDVI values in each bin). Some epochs were comprised of data from an adjacent year because of insufficient least cloud coverage, providing the acquisition date of the additional scenes corresponded with the March dry season. For example, the image epoch created for 1986 was made up of 2 scenes from 1986 and another from 1987, using the maximum value compositing approach (Holben, 1986) created initially for Advanced Very High-Resolution Radiometer (AVHRR) data. In this approach the greenest pixel of all the inputs determines the source pixel values that are used in the epochal image.

Sampling Approaches

Two sampling approaches were initiated to create the data sets used to understand the percent greenness changes across time and space (Figure 2.2). A combined county-protected area (CPA) variant summarized by polygon boundary file and a wall-to-wall variant spatially sampled the epochal layers by map-projected coordinate points. Both approaches were applied inside and in an approximately 20km buffer region outside of Trifinio Region. The distance of the outer buffer was related to our high-resolution WV2 imagery which extended to almost 20km beyond the boundaries of the region (Figure 2.3).

Figure 2.3. Trifinio Region Countries, Counties, and PAs, surrounded by buffer.

Point sampling applied variable sampling rates that were set at 10% and 50% depending on the PA extent (which in some cases had also been reduced by the presence of clouds, shadow, and water). PAs <50ha were sampled at a 50% rate, while all others were sampled at a 10% sampling rate, yielding more than 27,000 samples within PAs inside Trifinio, and almost 40,000 samples in PAs outside of Trifinio. The area outside of PAs was sampled at a 10% rate which yielded a total of about 722 thousand NDVI samples. Outside of the Trifinio Region sampling yielded a total of about 993 thousand samples. Sampling was carried out in TerrSet using the SAMPLE executable with vector and raster outputs; Boolean masks were created individually for each PA using the TerrSet BREAKOUT executable due to the need to apply variable sampling rates. All samples were converted using the TerrSet XYZIDRIS executable to a worksheet format. As a result, all samples, both inside and outside of Trifinio, contained their respective projected XY coordinates, and were given an identification (ID) number, and IDs to enable our analyses. Additional IDs included country identification codes (1, 2, and 3) for Guatemala, Honduras, and El Salvador, respectively, county name and source database code, code and name of nearest PA, a dummy variable to

indicate if sample was located within a PA, Euclidean distance from the sample to the nearest PA boundary in kilometers, area in kilometers of the nearest PA, greenness values in percent for each year of study, as well as greenness category number (1-5).

The second sampling approach, CPA, gathered knowledge from the same area, though the data values characterized the extent of a combination of the overlay of 56 PAs on 141 counties, in which 45 counties were within the Trifinio Region and 96 counties were located in the buffer outside of the region. Of the PAs, 48 are located within Trifinio, while six are located outside of the region in the 20km buffer. The labels of the 141-county polygons were multiplied by a factor of 100 and then the label values of the protected area polygons were added to the county labels. This allowed us to better manage the sources of PAs during analyses, as a single PA is often present in several counties, and in one case, in three different countries.

Basic Statistics

A total of 244 polygons comprised the administrative structures characterized in this study. Six PAs concerned solely with lakes were omitted at the outset as the study is about LCLUC. The portions of the 141 counties with and without 56 PAs across the combined inside of the Trifinio Region (45 counties) and outside (96 counties) were summarized by mode and mean values using a raster polygon mask for each combined CPA feature label (that had itself been masked by the 30-year cloud, shadow, and water features) (Figure 2.4). All summary operations used the TerrSet EXTRACT executable.

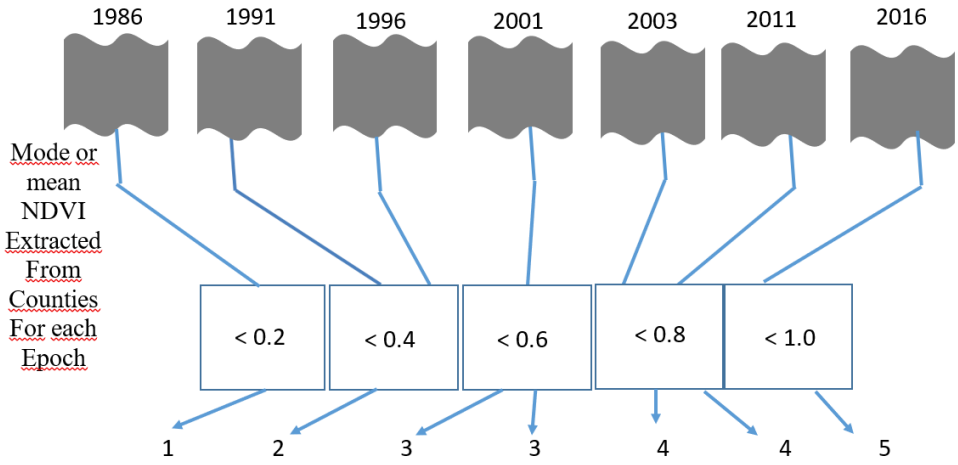


Figure 2.4. Data processing workflow for CPA group.

Figure 2.4 Mode and mean greenness values were extracted from each of the seven epochs with a county mask and reclassed to one of five 20% categories (called “quintiles” in this study). The quintiles were concatenated to each other to create short time series sequences and analyzed for transition direction.

We calculated a one-way analysis of variance (ANOVA) in R (version 3.3.3, 2017, “Another Canoe”, R Core Team 2017) with % greenness as the dependent variable and Year of % greenness as the independent variable to compare the means collected within each epoch to check on consistency. Additionally, we compared our greenness values with MODIS MOD13Q1 NDVI data to make sure our processing was not compromised by a misstep.

Cross-tabulation

We cross-tabulated greenness categories to quantify their change over time. Mean and modal values were extracted for each combination of county & PA for the three countries inside and outside of the Trifinio Region and these values were reclassed into 5 equidistant, fixed-boundary categories corresponding to 20% of spectral response for each epoch. Each category label was concatenated with the other labels corresponding to that county-PA as strings of values to help facilitate understanding of the spectral changes seen over the study period. For example, 2222345 is a CPA sample that started in greenness category 2 in 1986, stayed at the same level of greenness in the subsequent three epochs and then increased 1 category each in 2003, 2011, and 2016.

Clustering

Next, we clustered the data set to divide up its variability into a number of regions that might be explained. Clustering can occur spectrally or spatially, and we carried out both methods. Many clustering algorithms required a user-defined number of initial or maximum clusters. The number of relevant clusters in the percent greenness data was assessed by the ‘NbClust’ package in R (Charrad et al, 2014), which suggests the use of a majority rule to determine the outcome of 30 implemented indices. The number of clusters is useful to extract geospatial patterns that can be used to understand the mechanism and degree of land cover change in play during the dry season period used by the study. Twenty-eight of the indices are numeric. Two indices are graphical indicators (titled Hubert (Charrad et al, 2014) and Dindex

(Lebart et al. 2000)) that are interpreted by a visually determined “significant knee” in the graphical plots (Figure 2.5).

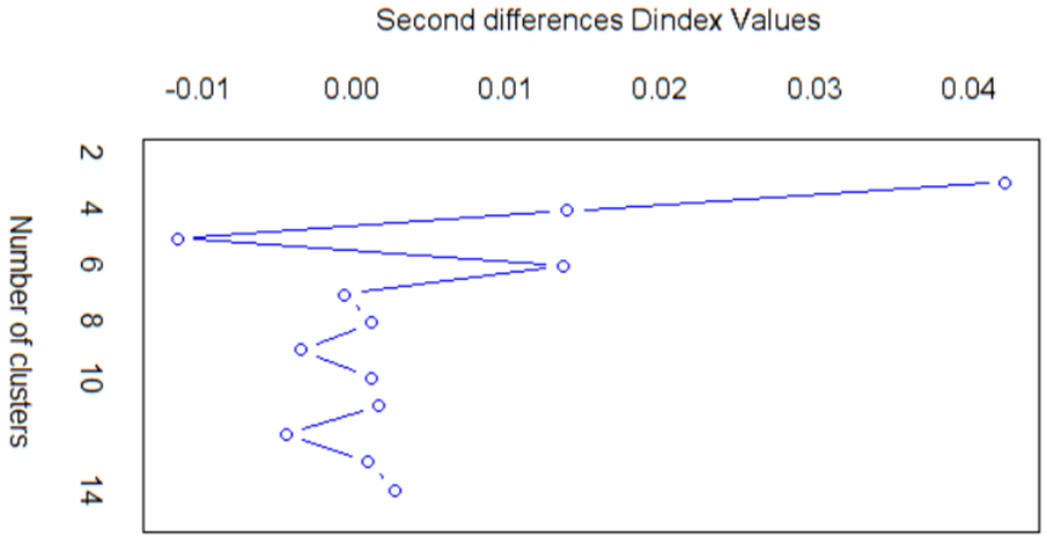


Figure 2.5. The Second differences Dindex Values graphic estimates 3 clusters of average percent greenness.

Using the predicted number of clusters (determined by the majority vote of the 30 indices), we predicted geographic cluster locations using the CLUSTER executable in TerrSet and also by using simple Microsoft Excel conditional statements to assess potential increasing or decreasing transition directions, discussed here as “slopes” (increasing, decreasing, or flat) and temporal transition length (e.g. for how many timesteps did a greenness transition continue). Consecutive higher or lower categorical placements were presumed to signify growth and loss of vegetation; two consecutive changes in the same direction higher or lower or two as long as the change is by two categories (Figure 2.6).

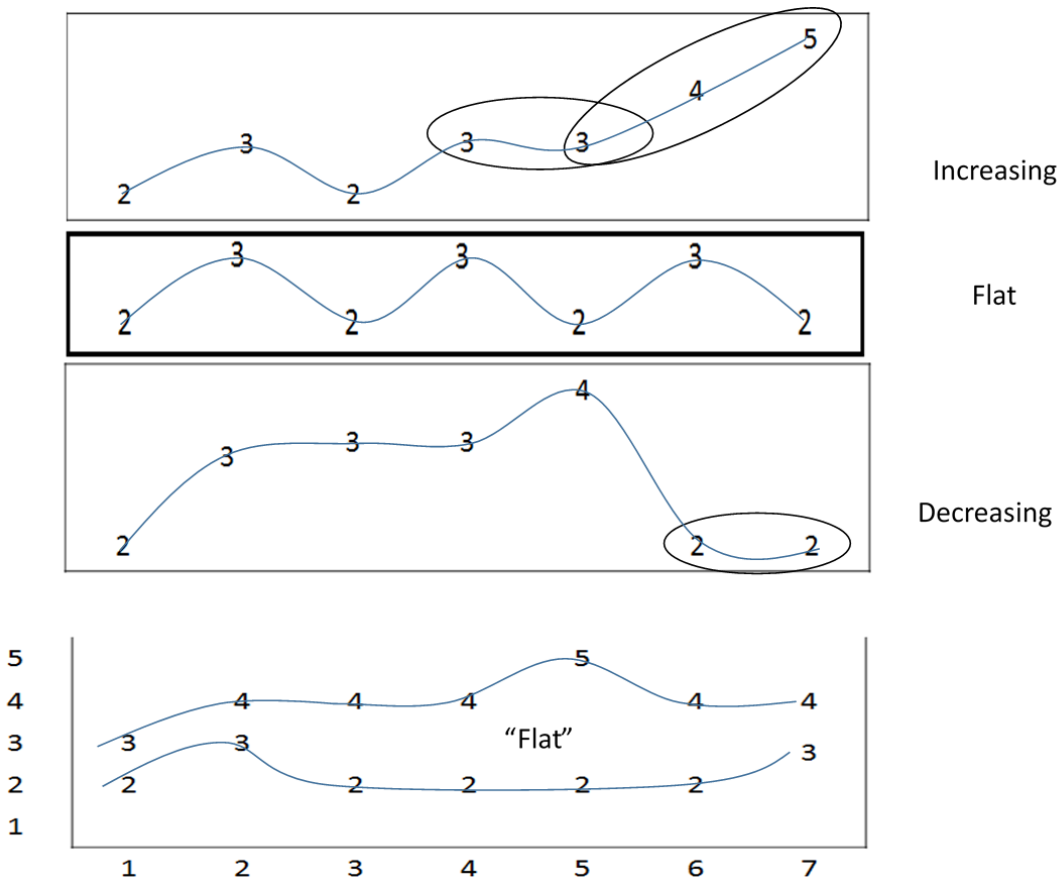


Figure 2.6. 30-year (7 epochs) assignments of slopes.

Short Time Series analyses

Short time series (STS) analytic software were used to indicate graphically the general shapes of land cover transition curves, and to visually sample locations of increasing and decreasing trend status that had been verified with WV2 imagery. STS using Cluster Analysis of Gene Expression Dynamics (CAGED) (Ramoni et al, 2002) and Short Time-series Expression Miner (STEM) (Ernst and Bar-Joseph, 2006) are Java-based tools have been used for more than a decade to express human genome profile data in sequences of eight-time points or fewer. Therefore, these tools allowed us to quickly visualize patterns of our series of seven-time points. STS profile plotting of the greenness samples using the gene expression tools proved immensely useful (especially CAGED) to quickly import, plot, and understand time series differences (to quickly see which sample or set of samples is at what potential growth point at which point in the series).

Modeling

We used R packages to study greenness trends over time using linear, nonlinear, and general linear model mixed effects models (applying the lm and nlme (R stats version 3.5, (R Core Team, 2017)) and R lme4 packages (Bates et al, 2017)). We assessed the statistical influence of administrative structures (i.e. county, protected area, country, and transboundary region) on greenness using InfoStat v. 2017d.

Normality assessment

Evidence of normality and significant differences between groups were assessed using R (version 3.3.3) (The R Foundation for Statistical Computing, 2017). To try to align the Landsat 7 data values more closely to Landsat 8 values, a calibration equation (Li et al. 2010; Roy et al, 2016) was applied to the 2003 data value.

Trends

Keeping in mind that the percent greenness data had been grouped in 5 categories by percent of positive NDVI, value changes across the seven epochs (using 4 questions) produced an index of the changing greenness data (that ranged at most from +3 to -3) permitting a comparison of temporal NDVI clustering:

To produce a slope index, a conditional statement was applied in MS Excel to assess incremental changes. This statement asked if the sum of the differences between temporally adjacent epochs (e.g. between Epoch 4 and Epoch 3, Epoch 6 and Epoch 5, and Epoch 7 and Epoch 6) is less than -1 then the slope is decreasing, and in other cases, if the sum of the differences in these same temporally adjacent epochs is greater than +1 then the slope is increasing, and in all other cases the slope is flat.

Applying the conditional statement above to categories ranging from 1-5, the only way to yield a number greater than 1 or less than -1 is to participate in a constant increasing or decreasing trend during the seven-epoch series. An additional set of equations was added to compare and accommodate the latest four epoch dates (last 15 years) (Figure 2.7) with the first three dates to learn what had happened more recently (e.g. 2001-2016). The trend equation applied for recent years is the same as the previous equation, but the starting year (2016) is differenced with 2011, 2011 is differenced with 2003, and 2003 is differenced with

2001; and then the sum of these differences is compared to -1 and +1, as earlier, to gain a picture of the most recent trend. These conditional comparisons of binned categories could mathematically lead to an index range extending from -4 to +4.

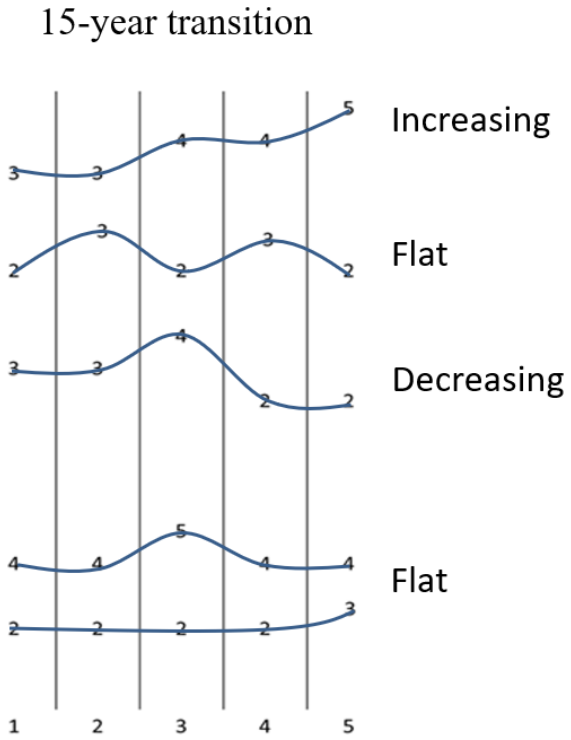


Figure 2.7. 15-year (5 epochs) assignment of slopes.

In long time series analyses it is common to smooth the series to eliminate/minimize spurious increases and decreases (Bayr et al, 2016), but that method does not help considerably in this case with only 7 temporal data points. Instead, we used a rule to allow the signal to rise or fall the width of one categorical bin, which happens frequently. Comparing summed differences between temporally-adjacent epochs greater than 1 and less than -1 accomplishes the same intent as filtering (without changing the values), as samples that rise and fall 2 or three times in a single sequence (e.g. 2323232) would be considered to have a flat slope.

Evaluation of trend directions (increasing, flat, decreasing) were easily plotted in CAGED. Individual plots of all 244 CPA average percent greenness timesteps showed the general trend directions graphically (Figure 2.8). Figure 2.8 A and Figure 2.8 B are indicative of flat trends, while the Figure 2.8 C is increasing. Flat trends can be located anywhere

vertically in the plots, but none were seen in the CPA approach outside of the 3-5 category (because the data are averaged).

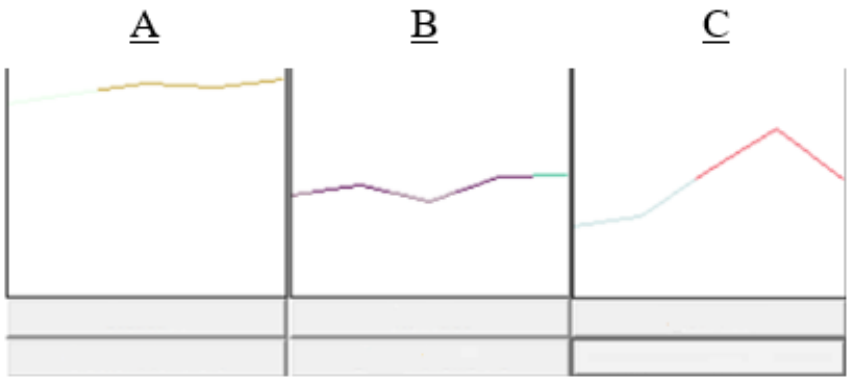


Figure 2.8. Sample flat (A & B) and increasing percent (C) greenness trends as plotted by CAGED.

Increasing trends (Figure 2.8. C) in this sample increase from 30% greenness in 1986 to 62% in 2011 and fall to 45% in 2016). The CPA data rise and fall in a subsequent epoch; none of the CPA data decrease to a point lower than a starting timestep.

When comparing frequency of transitions near PAs and proximity of transitions to PAs, an equation was applied to normalize the effect of the differences in frequency and rates due the concentric effect outside of the Trifinio Region. Given that the outside of Trifinio Region in out approximately 20km buffer is about 38% bigger than the area within the region, areas, sample frequencies, and therefore, rates were appropriately normalized by reducing those values with a scaling factor (Equation 2.1) before conversion to percent greenness trends:

$$0.7259 = 1/\ 10194\text{km2} / 7400\ \text{km2} \qquad \text{(Equation 2.1)}$$

At the most detailed level, each of the 722 thousand samples were evaluated for temporal and spectral trends and differences among the administrative structures. Proximity to nearest PA boundary was measured for each record using each sample's UTM coordinates. A set of 200 random "increasing slope" samples were verified as to their position on the landscape (in a field, on the edge of a field), interactively with WV2 data using ArcGIS taking into consideration that the sample source data were 30m resolution-sized Landsat pixels being

displayed on 2-meter resolution images. The distance from each field boundary was measured to each sample point. If the sample point was less than 30-meters distant from the start of the field, then the sample pixel was considered to be at the edge of a field, and if this distance was greater than 30 meters, then the sample was considered to be located within the field. Some fields were in use, while others appeared abandoned or were not in use. Some had green vegetation growing in them, but didn't appear to be in use.

Results

Basic Statistics

The minimum, maximum, average mean, and standard deviation of the different calculated percent greenness epochs (for Inside and Outside of the Trifinio Region) (Table 2.2). The average minimum and maximum percent greenness outside of the Trifinio Region is three and one percent higher, respectively. The average standard deviation in percent greenness is the same inside and outside. The highest maximum percent greenness values were outside of Trifinio Region, for four consecutive epochs. This contrasts with the average mean over the study period being two percent lower outside the region than inside.

Table 2.2. The minimum, maximum, average mean, and standard deviation for percent greenness epochs inside and outside of the Trifinio Region.

Inside Trifinio:					Outside Trifinio:			
Year	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev
1986	0.09	0.89	0.51	0.15	0.16	0.90	0.50	0.15
1991	0.14	0.91	0.52	0.16	0.16	0.95	0.49	0.17
1996	0.10	0.94	0.51	0.16	0.18	0.95	0.50	0.16
2001	0.10	0.91	0.52	0.16	0.18	0.95	0.50	0.17
2003	0.09	0.93	0.52	0.18	0.07	0.95	0.50	0.19
2011	0.17	0.93	0.56	0.15	0.12	0.89	0.56	0.15
2016	0.04	0.93	0.58	0.17	0.05	0.94	0.55	0.17

As noted, in one of our modeling runs, there had been a slight indication that year of data might be related with NDVI. An ANOVA comparing the means of epochs found that the independent variable Year was significant at the 0.001 level in relation to % greenness as

dependent variable. A subsequent Tukey HSD pairwise ad-hoc test was significant at the 0.05 level between the means of 2003, 2011, and 2016, and 2001 and earlier data. It was conceivable that the significant differences in mean greenness percent between 2016, 2011, and 2003, between Landsats 5, 7, and 8 was related to the image dates of those scenes (Table 2.1).

The relevant dates of the MOD13Q1 for Collection 6 were downloaded and imported into TerrSet to permit the percent greenness data values for the comparison with the imaging dates of the original Landsat images. The Landsat NDVI-converted to percent greenness compared favorably with a 500-pixel random sample extracted from the MODIS collection for same dates. There is no reason why these data should have had the same values (MODIS is collected by two different satellites, has very different resolution than Landsat (250m vs 30m), is collected using a different method, and our greenness epochs are not daily data), but they were similar enough to allow us gain confidence that the greenness data that we had sampled were not demonstrably different.

Cross-tabulation

Following the creation and cross-tabulation of categories, there were 18,556 groups of different strings, including 1111111, 2222222, 3333333, 4444444, and 5555555, containing 1-5 descriptive statistics about those temporal groups (though the distribution of those strings (Figure 2.9) is slightly skewed. Inspection of the variability showed that a cell in 1st category greenness bin in 1986 almost never reached category 5. In fact, only 1.7 thousandth of cases move from category 1 to category 5. There was never movement from category 1 to category 5 in five years and almost none in any ten-year period. Only about seven one hundredths of one-percent of all pixels in greenness bin 1 in 1986 regrew sufficiently to be included in greenness bin four or five by 2011 or 2016.

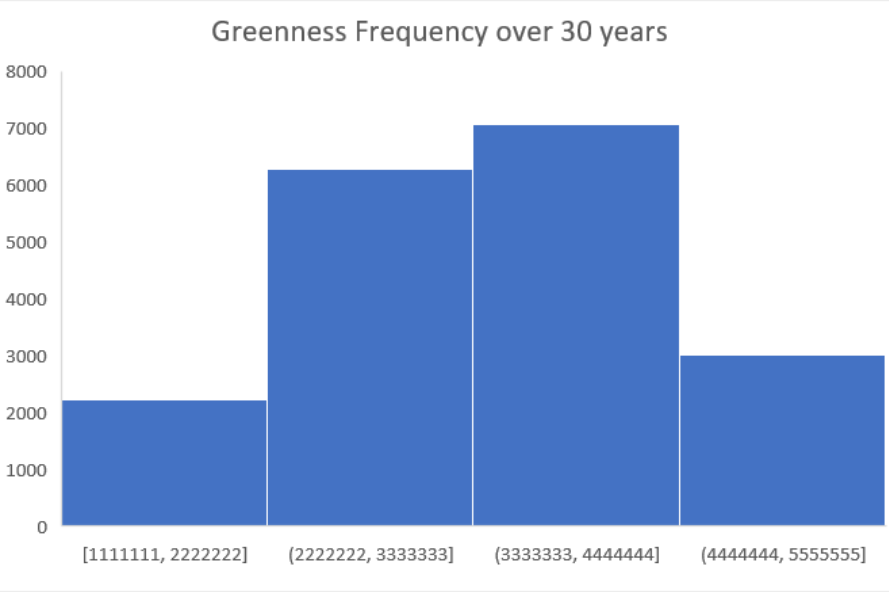


Figure 2.9. Greenness Frequencies in 30 Years organized in categories. There are no values less than 1111111, nor higher than 5555555.

Spatial autocorrelation is apparent in the source NDVI data (Moran's I was greater than .98 for all epochs analyzed).

Clustering

All indices of Nbclust were successful and revealed six different potential quantities of relevant clusters in the average percent greenness data (Table 2.3). The highest value in the Second Differences DIndex graphic determines the number of clusters in the data set (Figure 2.5). The majority rule indicated three clusters in the average percent greenness summary values, whereas the modal summary values proved to only have two clusters.

Table 2.3. Number of relevant average percent greenness clusters estimated by Nbclust.

Number of Relevant Clusters	Number of Reporting Indices
2	4
3	12
4	2
6	2
7	2
15	5

With three clusters identified corroborated by the NbClust analysis and CAGED plots, the sample data was separated using slope formulas in Microsoft Excel spreadsheets and subsequently verified using high-resolution WV2 images. Nearly 16% of the pixels over the entire study period inside and outside of the Trifinio Region did not demonstrate any change (as noted by the summed frequency of epochal bins that remained the same (1111111, 2222222, 3333333, 4444444, and 5555555, had 3,072,423 no change pixels out of a total of 19,324,574 pixels).

Modeling

No relationships, let alone significant relationships were found. Everything we tried, failed; these included linear models, logistic regressing, and generalized linear models. Normality testing ensued when it was discovered that none of the parametric approaches were working.

Normality

Graphical measures using box plots, qqplot plots, and histograms (Figure 2.10) offered evidence our data had not come from a normal distribution. Checks on skewness and kurtosis using the R package Nor.test (Gross & Liggs, 2015) “Tests for Normality” were carried out to confirm non-normality with the Anderson-Darling, Cramer-von Mises, Shapiro-Wilk, Lilliefors (Kolmogorov-Smirnov), and further follow-up with Pearson Chi Square tests, each of which significantly confirmed the dataset was not from a normal distribution (Gross & Liggs, 2015).

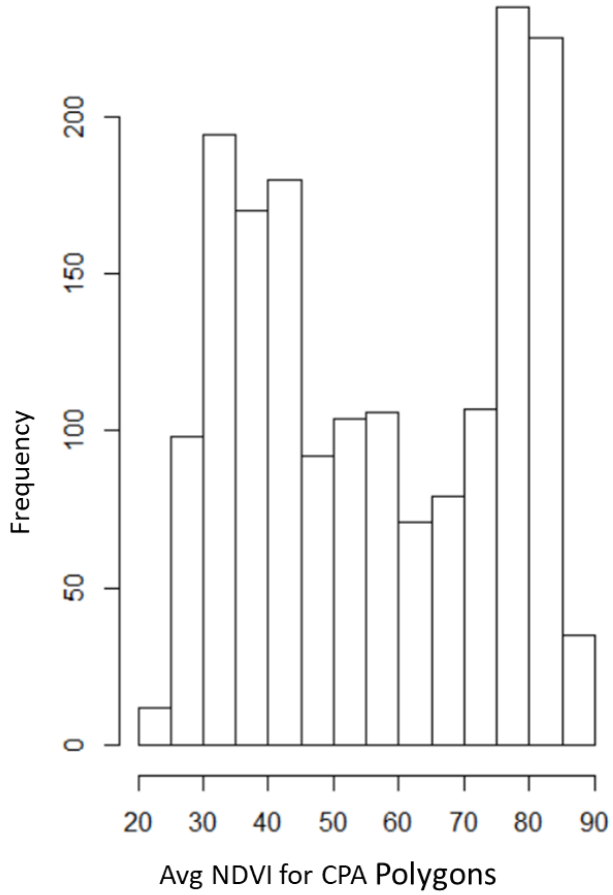


Figure 2.10. Multimodal histogram of average percent greenness data for combined CPA polygons across the Trifinio Region

Non-parametric significance testing was performed using Wilcoxon Rank Sum/Mann Whitney U tests in R (Table 2.4).

Table 2.4. Significant differences of greenness trends across administrative structures in the Trifinio Region.

Admin Structures	Located where?	Wilcoxon Rank Sum	Cohen's D	Effect Size	Outcome
County	All Inside vs All Outside	p-value 2.73e-06 ***	Est. -0.88	Large	Null Rejected
		95% CI [-0.08, -0.03]	95% CI [-1.3, -0.49]		
	All Inside vs Outside Avg of Means	p-value 1.29E-07 ***	Est. -0.27	Small	Null Rejected
		95% CI [0.41, 0.47]	95% CI [-63, -0.09]		
	All Inside vs Outside Avg of Modes	p-value 1.68E-06 ***	Est. -0.69	Medium	Null Rejected
		95% CI [-0.08, -0.01]	95% CI [-97, -0.43]		
Protected Areas	All PA vs non-PA Inside Trifinio	p-value 0.01892 **	Est. -0.56	Medium	Null Rejected
		95% CI [-0.19, -0.01]	95% CI [-.98, -0.14]		
	All Counties & PAs Inside vs Outside	p-value 0.01549 **	Est. 0.07	Negligible	Not Rejected
		95% CI [-0.08, -0.01]	95% CI [-.19, 0.32]		
	All inside PAs vs Outside PAs	p-value 4.65E-05 ***	Est. -0.83	Large	Null Rejected
		95% CI [-0.21, -0.07]	95% CI [-1.2, -0.42]		
	Trinational PA vs Others	p-value 1.80E-07 ***	Est. -1.7	Large	Null Rejected
		95% CI [-0.26, -0.11]	95% CI [-2.5, -0.9]		
Countries	Guat vs Non Guat	p-value 0.001126 **	Est. 1.3	Large	Null Rejected
		95% CI [0.07, 0.35]	95% CI [-.55, 2.0]		
	Hond vs Non Hond	p-value 0.02952 **	Est. -0.76	Medium	Null Rejected
		95% CI [-0.21, -0.007]	95% CI [-1.4, -0.15]		
	Elsa vs All Other PAs	p-value 0.6048	Est. 0.13	Negligible	INVALID
		95% CI [-0.09, 0.05]	95% CI [-.52 -0.77]		
	Elsa vs Guat/Hond	p-value 0.4074	Est. 0.29	Small	Not Rejected
		95% CI [0.12, 0.03]	95% CI [-.19, -0.78]		
	Guat vs Elsa/Hond	p-value 1.12E-06 ***	Est. 1.3	Large	Null Rejected
		95% CI [0.12, 0.34]	95% CI [0.8, 1.8]		
	Hond vs Elsa/Guat	p-value 1.78E-04 ***	Est. -0.77	Medium	Null Rejected
		95% CI [-0.21, -0.05]	95% CI [-1.2, -0.34]		
	Elsa vs Guat	p-value 0.3187	Est. 0.18	Negligible	Not Rejected
		95% CI [-0.09, 0.03]	95% CI [-.41, -.76]		
	Elsa vs Hond	p-value 5.52E-05 ***	Est. -1.0	Large	Null Rejected
		95% CI [0.07, 0.18]	95% CI [-1.5, -0.45]		
	Guat vs Hond	p-value 1.15E-05 ***	Est. -1.1	Large	Null Rejected
		95% CI [0.11, 0.22]	95% CI [-1.6, -0.57]		

*$p<0.05$ **; $p<0.001$ ****

Trends

The response to the first question on trends was broken down into three parts (to assess if there is statistically significant variation between the sample estimates organized by administrative structure), and reported in three tables. Table 2.4 presents the outcomes of non-parametric significance testing of greenness between counties, PAs, and countries inside and outside of the Trifinio Region using the Wilcoxon Rank Sum / Mann Whitney U (WMW) Test in R. The statistics include: the WMW p-values, the 95% confidence interval for the

WMW p-values, the Cohen's D estimate (in units of standard deviation) which evaluates the effect size of the WMW p-values, the 95% confidence interval for the Cohen's D estimate, the effect size, and outcome of the Null Hypothesis. The Null Hypothesis in this case is that there is no difference between the two greenness quantities in a particular administrative structure. The WMW p-value only tells us that there is or isn't a statistically significant difference between the greenness at the locations tested. However, Cohen's D will assess the standardized effect size (by differencing the two values and dividing by the standard deviation). If Cohen's D is negative, then the 2nd mean is larger than the first. In the first row, for example, all greenness values (organized by counties) inside and outside of Trifinio are compared.

Table 2.5 presents percent greenness trend differences geographically and across the administrative structures in units of percent greenness. Administrative structures are self-explanatory. Inside and outside of Trifinio are considered as a co-variable against which to measure county and PAs. Inside and outside refer to the Trifinio Region. A 95% confidence interval is offered for each WMW tested p-value. Cohen's D assesses the importance of the p-value and its effect size; and, Table 2.6. presents trends in percent like Table 2.5 (except the period evaluated is the most recent 15 years).

Table 2.5. Greenness trends across administrative structures over 30 years (in percent).

	30-Year Study Period					
Administrative Structures	Inside Trifinio			Outside of Trifinio		
	Increasing	Flat	Decreasing	Increasing	Flat	Decreasing
Counties	4.4	95.2	0.3	5	93	2
National – Guatemala	34.4	45.9	59.6	28	44.3	42.2
National – Honduras	56.8	39.2	31.0	49.7	31.7	18.6
National – El Salvador	8.8	14.8	9.4	22.4	24.0	39.7

Table 2.6. Greenness trends across administrative structures over latest 15 years (in percent).

	Latest 15-Year Trend					
Administrative Structures	Inside Trifinio			Outside of Trifinio		
	Increasing	Flat	Decreasing	Increasing	Flat	Decreasing
Counties	1.6	98.0	0.3	4	95	1
National – Guatemala	32.5	45.7	48.5	31.7	44.0	42.6
National – Honduras	40.7	39.9	48.2	31.5	32.5	29.0
National – El Salvador	26.8	14.4	3.3	37.1	23.5	28.4

The proximity of sources of increasing greenness suggests a relation to PAs (Table 2.1B), as the frequency of samples with increasing trends within Trifinio and near or within PAs is 35 times greater than the frequency of normalized samples outside of the region.

Table 2.7. Frequency of samples with increasing trends within and near PAs located both inside and outside of Trifinio.

Administrative Structures	Inside Trifinio 721166	Out of Trifinio 993427	Out of Trifinio Normalized
County	4.4% (32003) of 721K Samples Had Increasing Trends	5.5% (52081) of 993K Samples Had Increasing Trends	3.8% (37807) of 993K Norm. Samples
Increasing Trends in Prot Areas	3.8% (27564) Samples were found within PAs	1077 Samples were found within PAs	0.0007% (782) were found within PAs
Increasing Trend Proximity to Prot Areas	1924 < 1km 1714 >1<2km 1510 >2<3km 1514 >3<4km	1941 < 1km 2125 >1<2km 2370 >2<3km 2601 >3<4km	1409 < 1km 1543 >1<2km 1720 >2<3km 1888 >3<4km
National – Guat (incrsg)	11020	14572	10577
National -Hond (incrsg)	18183	25864	18775
National – El Sal (incrsg)	2800	11645	8453

Regrowth and Resurgence

An overlay of 3000 randomly selected with estimated increasing slope samples on the three most recent land use maps (one each for Guatemala, Honduras, and El Salvador, all created with RapidEye satellite images at 5x5m spatial resolution in the 2011-2014 period) only fell on coffee polygons about 1% of the time. The points fell adjacent to rivers, routes, in pastures, and on the edges of pastures. Some points fell on annual agricultural crops some on coniferous forest classes, some on secondary forest. In Guatemala, most fell on annual agriculture and brush.

Resurgence in abandoned fields and pastures

A total of 500 greenness samples with estimated increasing slope were randomly selected from the greenness databases sites of database samples were randomly selected. A total of 250 samples inside and 250 samples outside of the Trifinio Region were examined to check their location with the 2-meter WV2 images. Inside of the Trifinio Region, 104 increasing sites occurred in recent and abandoned agricultural fields, while 96 increasing samples fell on the edges of those fields. Outside of the Trifinio Region, 135 increasing sites occurred in recent and abandoned agricultural fields, while 65 samples fell on the edges of fields.

Discussion

With few exceptions, the data support significant differences in greenness transitions inside and outside of the Trifinio Region across the various administrative structures for the study period. Strangely, we would have expected the opposite findings for the null hypothesis' non-rejected in the cases of El Salvador vs Guatemala, El Salvador vs Guatemala and Honduras, because El Salvador has a much smaller amount of suitable land than other countries.

Significance testing indicated that greenness levels vary with significance across the different administrative structures, but not everywhere. More testing needs to be done to ascertain why. The significance of this is likely a mixed picture; conservation and natural resources management is likely not as centralized as supposed. The documents and papers describing Trifinio would have us believe that all is in common. But the outcome of this study

doesn't offer sufficient evidence for that. This overall finding might not have been expected under the assumption that common and centralized management would lead to more similarities.

The null hypothesis was that each of the two groups tested had equal medians, yet it has been shown that these data offered evidence that the medians were not the same and that the percent greenness trend values varied across all groups except El Salvador where noted.

The rates of change in trends inside and outside of Trifinio Region in the entire study period grouped by counties are the same with slightly differing magnitudes. There are differing (more increasing and decreasing) trends in Guatemala and Honduras vs El Salvador, and more flat trends in El Salvador when compared to Guatemala and Honduras. This last finding is potentially congruent with the conclusion by Hecht and Saatchi (2007) and Castaneda (2009) that boosted incomes from transnational remittances was impacting on-farm activity levels in El Salvador (though they did not study the Trifinio Region in their work). It is important to recognize 3 points that flat greenness trend does not necessarily mean a lack of activity. As the CAGED plots show, there are two important flat trends (one at the high level -- which could be equivalent to forests and agroforestry remaining strong and the other in the midlevel, and one at the mid-level that could be equivalent to non-forest dry season agriculture remaining constant.

Percent greenness trends in the latest 15-year period look somewhat similar to the 30-year trend, although the magnitudes of increasing trend have reduced, and decreasing have augmented in Honduras. Outside of Trifinio the magnitudes of decreasing greenness trend have changed for Honduras (increased by 50%) and El Salvador (decreased by 30%). Trends are largely the same for Guatemala in both periods.

The most likely reason for the significant differences in mean greenness percent between 2016, 2011, and 2003, between Landsats 5, 7, and 8 was related to the image dates of those scenes (Table 2.1). As historical rainfall was recorded in the range of dates in question (between the 9 and 21st of April), it is conceivable that our dataset could have recorded green up that had not occurred by that time in other epochs. We tried very hard to make sure the images were clustered around the same period of the year and had the least amount of cloud. The dry season is truly a very difficult time to assess green vegetation because many broadleaved species will have dropped their leaves, reducing the vegetation signal intended to

show variation in LCLUC trends; however, the current analysis would not have been possible in any other season with these data. Other studies in the region have had to accept the same set of circumstances (Nagendra et al, 2003; Southworth and Tucker, 2001). This has limited satellite-based investigations to coniferous vegetation types, such as the pine species reforested near the Celaque National Park in Honduras in the late 1990s.

The spatial autocorrelation result is an expected outcome in satellite remote sensing involving ecological and vegetation analyses, as it is quite logical that we would find similar features and species in geographic pockets with other similar variables. This quality can be important when analyses assess adjacency or proximity of pixel values to each other, such as in predictive spatial modelling. However, it is likely that none of the analyses undertaken in the CPA approach of this study could have been impacted by spatial autocorrelation because no adjacency or proximity-based measure of NDVI values was performed, and the greenness data associated with the county and PA polygons were elicited by modal or mean extractions using the polygon as a raster mask. We did examine the proximity of increasing slope greenness to PAs. Wulder and Boots (1998) prepared an extensive examination of the impact of spatial autocorrelation in remotely-sensed image analysis projects.

In any case, with respect to the proximity of increasing trends near PAs, almost no samples with increasing trends were found outside of the Trifinio Region; however, there are only 6 PAs outside of Trifinio compared with 50 PAs inside the region. While only 4.4% of samples associated with increasing percent greenness trends were found within PAs inside Trifinio, this amount of 35x greater inside the regional PAs than those outside the region suggests a need for follow up work to find out why. An interesting trend direction was noted in relation to proximity to PAs which proves the hypothesis to be correct near the PAs; Schlesinger et al. (2017) found this same trend in its examination of forested LCLUC surrounding the Montecristo Trifinio Trinational Park; that is, that frequency of forest is higher adjacent to the park and dissipates moving away from the park, whereas outside of Trifinio the frequency of users is different. The frequency of users is smaller near the PAs and greater away from them.

An examination of the five categorical greenness bins measured at the pixel scale summed by county level offers no evidence to show that previously degraded lands (of the first greenness category) in the Trifinio Region are part of any large scale regional recovery or

transition to a stage congruent with agroforestry or forest cover. No large greenness transitions were estimated. These county data showed that only 0.07% of pixels were found to have recovered from category 1 greenness. It is conceivable that land deforested for replanting to agroforestry (e.g. coffee) could have been converted after the dry season in a single epoch study date, and might have been measured as woodland, bush, or forest cover and then as agroforestry in a subsequent epoch. But given the very small percentage of pixels recovering to any extent in the region, it is more than likely that Trifinio Region is not participating in any resurgence reported by other studies.

There is little evidence that this is occurring because of the size of the increases and the locations of the increasing growth pixels, occurring within fields or on the edges of fields. Whatever they are, they are not going to replace or act as replacement forest. Coffee shading is problematic for agroforestry identification by eye and by remote sensing, but no overall resurgence of greenness was found using the methods applied in this study.

Only a small number of researchers have been working on land cover change near the Trifinio Region and most during the 2000-2012 period with differing types of data reported and varying amounts and types of forest cover regrowth that occurred in the latter part of the 1990s. Redo et al (2012) reported conifer regrowth, but no broadleaf recovery, and net deforestation in Guatemala. Vaca (2012) reported recovery as complex and resulting from passive properties in Southern Mexico at 0.2% per year, with Chowdhury et al (2004), and Turner (2010) reporting similarly. Bray (2009) reports Guatemala forest recovery nationally at 3% annually, with forest plantations of 133,000 ha versus coffee agroforestry at 260,000 ha, but he did not work in the Trifinio Region. The forest plantations in El Salvador are very visible on satellite images, but their dimensions did not appear to change over the study period. Southworth et al. (2004) worked in the Celaque National Park (Honduras) and reported of 25% land cover change occurring there before 2000; none of this resurgence was visible through this work.

There is limited evidence of increasing greenness occurring more in abandoned fields and pasture lands. Only about 50% of the 200 random samples were found in recent fields or abandoned pastures inside of Trifinio (Table W), and about 68% were found outside of the Trifinio Region. Abandoned fields and pasture lands are a regular feature of the Trifinio Region landscape, but it is not very easy to recognize them in the dry season, even with very

high-resolution satellite images. It is apparent that greenness trends are occurring passively near rivers, transport routes, and on the edges of areas of heavier land use (e.g. pasture, agriculture).

As noted above, if vegetation in greenness category 1 (bare soil degraded grasslands) remains; it does not recover to a level sufficient to support agriculture or agroforestry within 25-30 years. If vegetation in greenness category 2 becomes extremely degraded and moves to category 1, it will likely not recover enough to become a member of category 4 or 5, if these data prove that history of these samples is any predictor of future behavior. This does not bode well for the future of degraded abandoned pastures and other fields in the Trifinio Region and beyond.

A review of preprocessing methods did not find any difference in data handling, and following a rerun of data preparation from source materials, the ANOVA still proved significant for the years in question. A follow-up investigative check of historical rainfall for the periods of March 16- April 21 for both 2011 and 2016 showed that rainfall in millimeter quantities were recorded at the Esquipulas Airport during those days. The historical data show that 11.5 mm of rain fell during that period in 2011, and 5 mm fell during the same period in 2016. Those amounts may seem very small, but in the years in question they amounted to about 5% of annual rainfall for those locations.

A literature search into very short time series approaches yielded no suitable models. Much of the relevant literature about timeseries of NDVI use data from 250-meter MODIS MOD13Q1, with as many as 23 images annually with 16-day revisit frequency (Redo, 2012) (or are using 1km resolution SPOT VEGETATION data at an annual sequence (Tucker et al, 2005). However, because the Trifinio Region is so cloudy, most of the time, when a Landsat satellite passes by 50% of the region is covered by cloud cover. Even with the max greenness method we used, we still lost nearly 10% of the inside of the Trifinio Region as it as was always covered by clouds. Like in many other tropical locales, it is not possible to assess vegetation with optical sensors outside of the “dry” seasons. Indeed, the cloudiness observed made more temporally frequent collection of imagery impossible. The Landsat data collected are the best we could amass for this region without merging additional sensors (e.g. MODIS) of more frequent coverage, but with considerably larger pixels. Analyses of available Landsat data only permitted the creation of seven epochs over the 30 years study period. It is

conceivable that in time with regular imaging by the Sentinel 2A and 2B optical sensors, there may be additional coverage as the overpass times (10-day revisit frequency in Central America) will be more frequent than Landsat.

Land cover change evaluation by time series is commonly carried out with longer or more complete series of annual or seasonal NDVI data; the data available through passive sensing in this part of the world is replete with cloud cover for much the year. There is only a short period of time with medium resolution images to be able to view the land surface due to clouds and shadows and long revisit times. MODIS data, widely used and available since 1999, is commonly evaluated, however while useful for this work to corroborate its NDVI values, the minimum pixel size (250m) is counterproductive for measuring in tropical Latin American tropical environments dominated by small farmers working small fields near and undercover of shaded agroforestry canopies (near coffee, cocoa, and others). Even Landsat's 30-meter pixel resolution is difficult to use for analyses of most small farmer holdings. The Worldview 2 data, on the other hand, with higher-resolution and more spectral depth is certainly sufficient (when plants are young to count the individual plants in planted rows), but these too can have limited availability because as a commercial sensor it is not turned on unless there is a paying client, and this limits its availability in archival use due to more potential for cloud cover given reduced coverage.

Uncertainty

Potential uncertainties of this work including the timing of the images, signal saturation, lack of correction for slope and aspect factors that influence illumination differences, and stand-age. It is supposed that much broad-leaved vegetation was in a leaf-off period for the extent of the study, thus the timing of the images centered around the month of March and in one case April may have influenced the study's outcome; however, no other image dates were available to try to assess the cover change.

To reduce uncertainty variation at the project outset, all the data used in the study were surface reflectance-corrected Eros Science Processing Architecture (ESPA) Level-2 science products. No caveats nor constraints associated with those products were triggered by our image acquisition dates. All pixels remaining with cloud cover or shadow after the maximum-value compositing effort to build the seven epochs were removed from the dataset, and all the

cloud, shadow, and water pixels in the composite 30-year mask were also removed from all querying GIS layers created for each of the administrative structures. The NDVI values estimated for the epochs of the 2001-2016 epochs compared favorably for a 500-sample set of random locations with MODIS Mod13Q1 Collection 6 NDVI data.

Yet is possible that external influences may have affected the variability of response in the pine-oak dominated region surrounding Celaque National Park, for example, as the park is situated on an upland on the edge of the Trifinio Region. There are considerable slope and aspect differences extending outward in all directions (with a mean slope of 30%) from the center of the park to its boundaries on all sides. However, no differences in apparent greenness were more or less visible at any location around the park.

Moderate to high-density biomass has been shown to impact NDVI saturation (Viña et al, 2004), and it is conceivable that this impacted our understanding of change in the Celaque region. Some of those pines reportedly grew back (the 'old and permanent forest regrowth' class) and achieved 25% canopy closure in just 9 years from previously farmed fallows (Nagendra et al, 2003) in response to a countywide logging ban initiated in 1987. Some others, two classes would have had to achieve a high level of biomass in just five years as they are noted as having not been a forest in 1991 and being so in 1996. While possible, it is questionable that these trees could have achieved a moderate to high biomass in such a short period of regrowth.

Conclusion

Greenness is not static among most counties, PAs, and countries of the Trifinio Region. Land cover and land use is always changing due to local needs for agricultural, agroforestry, and pastoral uses. Many lands prepared for annual cropping are burned to eliminate stubble and litter, and while this is evident in high-resolution images, it was not recognizable on our medium resolution Landsat images. A tremendous amount of the lands that make up the Trifinio Region (both inside and outside) vary in greenness percentages, but were not seen to be changing their greenness quantities temporally to any large degree plus or minus. The spectral slope was largely measured to remain flat (with little to no change). Among those that remain flat, there are many that increase and decrease every five-years,

fluctuation in direction more than a few percentage points each epoch. Almost no extreme events in one direction or another were found.

Only a very small percentage of lands might be experiencing regrowth or resurgence of green cover. However, there was no clear signal of any location that is experiencing regrowth more than any other, contrary to reports by regional researchers; all signs of increasing greenness were sporadic. Surprisingly, there was no sign of the coniferous resurgence reported in the vicinity of the Celaque National Park of Honduras (Redo et al, 2012; Nagendra et al, 2003). Some locations of increasing greenness can be found in fields, on the sides of fields, or as a factor in abandoned fields and pastures, but there was not any clear signal to favor one location over another.

In other tropical countries, Brazil and Peru, for example, squatters in agriculture and pastoral activities have taken up positions outside of PA boundaries and their clearing fires and transitions are easily found on satellite images, but nothing has yet been found like that in the Trifinio Region. Management efforts in sub-watersheds seem to be working, because no major greenness transitions were found in these areas more than others. Greenness appears on pasture edges, remnant patches of forests between agricultural areas, and in brush, but not in upland coffee landscapes. However, transitions are occurring in areas of heavier land use, following human transport routes, rivers, and highways, and likewise, vegetation is appearing in populated urban centers.

Interestingly, though, lands of very low greenness, perhaps degraded pastures, were almost never seen to recover. Once the landcover had been allowed to degrade to the level of the first quintile, the greenness quality never recovered. On this same note, about 5% of lands never change, remaining at the same level of greenness today as they were 30 years ago. These findings should offer some good news to regional natural resource managers, who can know that their efforts reduce forest resource extraction in areas of intact forest cover, especially the micro-watershed recharge areas, are seeing positive results. Additionally, degraded lands of continually low amounts of vegetation, may offer an opportunity for their rehabilitation.

Acknowledgements

We gratefully acknowledge the assistance from our land management partners in this research including the Tropical Agricultural Research and Higher Education Center (CATIE) 's Mesoamerican Agroenvironmental Program (MAP Norway) for their assistance in field access and data collection, Central American Commission for Environment and Development (CCAD), Trinational Commission of the Trifinio Plan (CTPT), Proyecto Trinacional Café Especial Sostenible (Protcafes), GIZ Forests and Water Program, and the US Fish and Wildlife Service.

Funding

This work was funded by NASA LCLUC Grant NNX13AC70G (K. Jones, PI) and MICITT/CONICIT in Costa Rica, and the University of Idaho Graduate Fellowship Program.

References

Aguilar, A. (2003). Patterns of forest regeneration in Celaque National Park, Honduras. *Online Journal of Space Communication*, *3*, 31.

Allnutt, T. F., Asner, G. P., Golden, C. D., & Powell, G. V. N. (2013). Mapping Recent Deforestation and Forest Disturbance in Northeastern Madagascar. *Tropical Conservation Science*, *6*(1), 1–15. https://doi.org/10.1177/194008291300600101

Argotty, F., Zamora Perreira, J. C., Brenes Perez, C., Schlesinger, P., Cifuentes, M., & Imbach, P. (2017, July). Insumos metodológicos para el establecimiento de niveles de referencia para REDD+: deforestación y recuperación de cobertura forestal en la Región Autónoma de la Costa Caribe Norte, Nicaragua. USAID. Retrieved from https://www.catie.ac.cr/que-es-catie/129-programas-integradores/341-programa-regional-de-cambio-climatico-de-usaid

Asner, G. P., Knapp, D. E., Balaji, A., & Paez-Acosta, G. (2009). Automated mapping of tropical deforestation and forest degradation: CLASlite. *Journal of Applied Remote Sensing*, *3*(1), 033543-033543-24. https://doi.org/10.1117/1.3223675

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *ArXiv:1406.5823 [Stat]*. Retrieved from http://arxiv.org/abs/1406.5823

Bayr, C., Gallaun, H., Kleb, U., Kornberger, B., Steinegger, M., & Winter, M. (2016). Satellite-based forest monitoring: spatial and temporal forecast of growing index and short-wave infrared band. *Geospatial Health; Vol 11, No 1 (2016): Valencia Issue*. Retrieved from http://www.geospatialhealth.net/index.php/gh/article/view/310/452

Bray, D. B. (2009). Forest Cover Dynamics and Forest Transitions in Mexico and Central America: Towards a “Great Restoration”? In H. Nagendra & J. Southworth (Eds.), *Reforesting Landscapes* (Vol. 10, pp. 85–120). Dordrecht: Springer Netherlands. Retrieved from http://www.springerlink.com/index/10.1007/978-1-4020-9656-3_5

Castaneda, H. (2009). *Analysis of the spatial dynamics and drivers of forest cover change in the Lempa River Basin of El Salvador*. [Gainesville, Fla.]: University of Florida. Retrieved from http://purl.fcla.edu/fcla/etd/UFE0024235

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014, April 13). Determining the Best Number of Clusters in a Data Set. Retrieved from https://cran.r-project.org/web/packages/NbClust/NbClust.pdf

Chowdhury, R. R. (2010). Differentiation and concordance in smallholder landuse strategies in southern Mexico's conservation frontier. *Proceedings of the National Academy of Sciences*, *107*(13), 5780–5785. https://doi.org/10.1073/pnas.0905892107

Deering, D. W. (1975). Measuring "forage production" of grazing units from Landsat MSS data. In *Tenth hnt. Symp. on Remote Sensing of Environment* (pp. 1169–1178). Ann Arbor: University of Michigan.

Eastman, J. R. (1989). IDRISI : A geographic information system for international development. Presented at the
Conference on Information Technologies for Developing Countries, California: University of Southern California.

Environmental Systems Research Institute (ESRI). (2014). *ArcGIS Release 10.5*. Redlands, CA.

Ernst, J., & Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics*, *7*(1), 1–11. https://doi.org/10.1186/1471-2105-7-191

Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley Jr, R. D., Beckmann, T., … Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment*, *194*, 379–390. https://doi.org/10.1016/j.rse.2017.03.026

Foucart, S. (2011, July 12). A race for land is destroying the Guatemalan rainforest. *The Guardian*. Retrieved from https://www.theguardian.com/environment/2011/jul/12/guatemala-rainforest-deforestation-farming-foucart

Hecht, S. (2010). The new rurality: Globalization, peasants and the paradoxes of landscapes. *Forest TransitionsWind Power Planning, Landscapes and Publics*, *27*(2), 161–169. https://doi.org/10.1016/j.landusepol.2009.08.010

Hecht, S. B., & Saatchi, S. S. (2007). Globalization and forest resurgence: Changes in forest cover in El Salvador. *Bioscience*, *57*(8), 663–672.

Heinrich, V., Krauss, G., Gotze, C., & Sandow, C. (2012). Index Database: A database for remote sensing indices. Retrieved September 23, 2016, from http://www.indexdatabase.de/info/credits.php

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, *7*(11), 1417–1434.

Jimenez, A. J. (2014). *Final Report of ES Forest Map 2011, July 2014*. San Salvador.

Kanniah, K. D., Najib, N. E. M., & Vu, T. T. (2016). FOREST COVER MAPPING IN ISKANDAR MALAYSIA USING SATELLITE DATA. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. XLII-4/W1). Kuala Lumpur, Malaysia: ISPRS. Retrieved from https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-4-W1/71/2016/isprs-archives-XLII-4-W1-71-2016.pdf

Lebart, L. (2000). Contiguity Analysis and Classification. In W. Gaul, O. Opitz, & M. Schader (Eds.), *Data Analysis* (pp. 233–243). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58250-9_19

Li, F., Jupp, D. L. B., Reddy, S., Lymburner, L., Mueller, N., Tan, P., & Islam, A. (2010). An Evaluation of the Use of Atmospheric and BRDF Correction to Standardize Landsat Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(3), 257–270. https://doi.org/10.1109/JSTARS.2010.2042281

MARN. (April 20016). MARN lanza el Plan Nacional de Restauración y Reforestación | MARN | Ministerio de Medio Ambiente y Recursos Naturales [Government]. Retrieved July 11, 2016, from http://www.marn.gob.sv/marn-lanza-el-plan-nacional-de-restauracion-y-reforestacion/

MASEK, J. G., VERMOTE, E. F., SALEOUS, N., WOLFE, R., HALL, F. G., HUEMMRICH, K. F., … LIM, T. K. (2013). LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2. https://doi.org/10.3334/ORNLDAAC/1146

Nagendra, H., Southworth, J., & Tucker, C. (2003). Accessibility as a determinant of landscape transformation in western Honduras: linking pattern and process. *Landscape Ecology*, *18*(2), 141–158.

OAS. (1993). Constitución Política de la República de Guatemala. Retrieved from https://www.oas.org/juridico/mla/sp/gtm/sp_gtm-int-text-const.pdf

Pacifici, F., Longbotham, N., & Emery, W. J. (2014). The Importance of Physical Quantities for the Analysis of Multitemporal and Multiangular Optical Very High Spatial

Resolution Images. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(10), 6241–6256. https://doi.org/10.1109/TGRS.2013.2295819

Ramoni, M. F., Sebastiani, P., & Kohane, I. S. (2002). Cluster analysis of gene expression dynamics. *Proceedings of the National Academy of Sciences*, *99*(14), 9121–9126. https://doi.org/10.1073/pnas.132656399

Redo, D. J., Grau, H. R., Aide, T. M., & Clark, M. L. (2012). Asymmetric forest transition driven by the interaction of socioeconomic development and environmental heterogeneity in Central America. *Proceedings of the National Academy of Sciences*, *109*(23), 8839–8844. https://doi.org/10.1073/pnas.1201664109

Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation* (Final Report) (p. 390). Greenbelt, MD: NASA GSFC.

Roy, D. P., Kovalskyy, V., Zhang, H. K., Vermote, E. F., Yan, L., Kumar, S. S., & Egorov, A. (2016). Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Landsat 8 Science Results*, *185*(Supplement C), 57–70. https://doi.org/10.1016/j.rse.2015.12.024

Sader, S. A., Hayes, D. J., Hepinstall, J. A., Coan, M., & Soza, C. (2001). Forest change monitoring of a remote biosphere reserve. *International Journal of Remote Sensing*, *22*(10), 1937–1950.

Sader, S. A., Roy Chowdhury, R., Schneider, L., & Turner, B. L. (2004). Forest change and human driving forces in Central America. Retrieved November 2, 2015, from http://geography.indiana.edu/faculty/roychowdhury/bookChapters/Sader,%20Roy%20Chowdhury%20et%20al%202004_Forest%20change%20and%20human%20driving%20forces%20in%20Central%20America.pdf

Schlesinger, P., Muñoz Brenes, C. L., Jones, K. W., & Vierling, L. A. (2017). The Trifinio Region: a case study of transboundary forest change in Central America. *Journal of Land Use Science*, *12*(1), 36–54. https://doi.org/10.1080/1747423X.2016.1261948

Southworth, J., & Tucker, C. (2001). The Influence of Accessibility, Local Institutions, and Socioeconomic Factors on Forest Cover Change in the Mountains of Western Honduras. *Mountain Research and Development*, *21*(3), 276–283. https://doi.org/10.1659/0276-4741(2001)021[0276:TIOALI]2.0.CO;2

Tarigan, S. D. (2016). Land Cover Change and its Impact on Flooding Frequency of Batanghari Watershed, Jambi Province, Indonesia. *Procedia Environmental Sciences*, *33*(Supplement C), 386–392. https://doi.org/10.1016/j.proenv.2016.03.089

The R Foundation for Statistical Computing. (n.d.). The R Computing Environment (Version 3.3.3) [R]. The R Foundation for Statistical Computing.

Tucker, C. J. (1978). *Red and Photographic Infrared Linear Combinations for Monitoring Vegetation* (NASA Technical Memorandum No. 79620) (p. 33). Greenbelt, MD: NASA GSFC. Retrieved from http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19780024582.pdf

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., … El Saleous, N. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, *26*(20), 4485–4498. https://doi.org/10.1080/01431160500168686

Tucker, C. M., Munroe, D. K., Nagendra, H., Southworth, J., & others. (2005). Comparative spatial analyses of forest conservation and change in Honduras and Guatemala. *Conservation and Society*, *3*(1), 174.

Turner, B. L. (2010). Sustainability and forest transitions in the southern Yucatán: The land architecture approach. *Land Use Policy*, *27*(2), 170–179. https://doi.org/10.1016/j.landusepol.2009.03.006

USAID. (2014). *VULNERABILITY AND RESILIENCE TO CLIMATE CHANGE IN WESTERN HONDURAS*. Retrieved from http://community.eldis.org/.5b9bfce3/Western%20Honduras%20VA_ENGLISH_CLEARED.pdf

USGS. (2014). *Landsat Surface Reflectance Climate Data Records*. Retrieved from https://pubs.usgs.gov/fs/2013/3117/pdf/fs2013-3117.pdf

USGS. (2017, October). Landsat 4-7 Surface Reflectance (LEDAPS) Product Guide. USGS. Retrieved from https://landsat.usgs.gov/sites/default/files/documents/ledaps_product_guide.pdf

Vaca, R. A., Golicher, D. J., Cayuela, L., Hewson, J., & Steininger, M. (2012). Evidence of Incipient Forest Transition in Southern Mexico. *PLoS ONE*, *7*(8), e42309. https://doi.org/10.1371/journal.pone.0042309

Viña, A., Henebry, G. M., & Gitelson, A. A. (2004). Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the wide dynamic range vegetation index. *Geophysical Research Letters, 31*(4), L04503. https://doi.org/10.1029/2003GL019034

Yengoh, G. T., Dent, D., Olsson, L., Tengberg, A. E., & Tucker, C. J. (2014). *The use of the Normalized Difference Vegetation Index (NDVI) to assess land degradation at multiple scales: a review of the current status, future trends, and practical considerations* (p. 80). Lund, Sweden: Lund University Centre for Sustainability Studies - LUCSUS. Retrieved from http://www.stapgef.org/stap/wp-content/uploads/2015/05/Final-report-The-use-of-NDVI-to-assess-land-degradation-G.-Yengoh-et-al..pdf

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment, 118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

CHAPTER 3: ADVANCES IN MAPPING COFFEE PLANTATIONS: THE CASE OF THE TRIFINIO REGION, CENTRAL AMERICA.

Abstract

Coffee (*Coffea* arabica) is one of the most important agricultural crops grown in tropical regions, with added implications for local ecosystems, biodiversity, and water management. Determining the location and extent of coffee plantations, especially those using shade production methods is therefore an important endeavor. However, mapping coffee agroforestry systems using remotely sensed images is challenging because of the similarity between the spectral characteristics of coffee, and native tropical leaves and woodland structures. This research applies texture analysis to Worldview 2, Landsat 8, Sentinel 1 & 2 images, assessing the accuracy of classification in the identification of coffee. We compared combinations of image bands and textures and the latter were not found to be significant predictors. Neural network, Naïve Bayes, K-Nearest Neighbor, Maximum Likelihood, Random Forest, and Support Vector Machines algorithms were applied in R with RStoolbox (Leutner & Horning, 2017) to identify pasture, forest, coffee classes of shade, sun, adult, and immature. Overall coffee map accuracy was analyzed with confusion matrices and kappa indices. Results indicated that a maximum of 85-87% accuracy is achievable with Landsat data alone; a hybrid optical, infrared, vegetation index with effective incidence angle predictors was shown to be effective; however, no increased accuracy over methods used by previous studies was attained.

Keywords: Coffee, machine-learning, remote sensing, shade, sun, texture, Trifinio

Resumen

El café (Coffea arabica) es uno de los cultivos agrícolas más importantes que se cultivan en regiones tropicales, pero también tiene implicaciones para el ecosistema local y la gestión del agua. La determinación de la ubicación y extensión de las plantaciones de café es, por lo tanto, un esfuerzo importante. Sin embargo, mapear las plantaciones de café usando imágenes obtenidas por teledetección es un desafío debido a la similitud entre las características espectrales del café y los árboles tropicales. Esta investigación se aplica al

análisis de textura de imágenes de Worldview 2, Landsat 8, Sentinel 1 y 2, evaluando la precisión de la clasificación en la identificación del café. Comparamos combinaciones de bandas de imagen y texturas y las últimas no se encontraron como predictores significativos. Los algoritmos de la red neuronal, Naïve Bayes, K-Nearest Neighbour, Maximum Likelihood, Random Forests y Support Vector Machines se aplicaron con R para identificar pasto, bosque y café, las clasesde sombra, sol, adulto e inmaduro. La precisión general del mapa del café fue analizado con matrices de confusión. Los resultados comentaron que casi el 93 por ciento de la precisión del mapeo es posible con un clasificador de redes neuronales y datos de Landsat 8 solo; se demostró que es eficaz un índice híbrido óptico, infrarrojo, de vegetación con predictores de ángulo de incidencia efectivo. Se alcanzó una mayor precisión sobre los métodos en estudios previos.

Palabras claves: Aprendizaje automático, teledetección, café con sombra, café en plein sol, textura, Trifinio

Introduction

The role of coffee in tropical land cover change

Coffee, a cash crop, is the most important agricultural commodity in the world and is only second to oil in terms of trade value (Cordero-Sancho and Sader (2007); Donald, 2004), it has played a complex role in land use-led land cover changes in rural subsistence farming in neotropical dry forest regions (McCook, 2017). While contributing considerably to regional deforestation growth with a nearly three-fold contribution from land use expansion (Schlesinger et al, 2017), its production types (full sun, and under shade) have made positive contributions to ecological services (Blackman et al, 2012). Shade coffee is involved with climate and nutrient regulation, controlling erosion, insect abundance, and providing habitat support for bird species involved in seed disbursals and pollinators (Jezeer et al, 2017; Jha et al, 2014; Classen A et al. (2014); Schmitt-Harsh et al (2013); Richards and Mendez, 2013; Blackman et al (2012); Rice, 2010; and Greenberg et al., 1997). Sun coffee contributes to some of these benefits creating higher yield due to higher planting density permitted with greater inputs (labor, irrigation, and chemicals) and shorter time to market (Donald, 2004); though, shaded coffee is much more effective (Rice, 2010). In El Salvador, despite the higher

yield opportunity, the great majority of farms (95%) pursue shade coffee production (Blackman et al, 2012). Understanding these benefits can lead to good management decisions and subsequently to improved crop production, employment, and sustainability.

The drivers of coffee land use changes are complicated and vary in that existing coffee development areas are necessarily not the same as those traditionally found near undeveloped native forests (Blackman et al, 2012) where development is usually related to flat, fertile land, near to roads and markets that could increase profits due to decrease access costs (Kleinschroth and Healey, 2017; Geist & Lambin, 2002; Geist & Lambin, 2001). It depends on the existing dominant cleared land use, posits Blackman et al (2012). In their modeled examination of El Salvador near two urban and one rural settings from 1990-2010, they found that proximity to urban areas did not explain clearing. They did find, however, in the rural setting that traditional drivers (subsistence agriculture, pasture, and livestock production provide greater returns) do drive coffee plot forest-clearings. Monitoring of these land use changes over time, especially those that propel coffee agroforestry has proven difficult, especially in areas where today there is currently little outright deforestation, and populations are being led to coexist in sustainable fashion, develop areas to protect biodiverse plant and animal species, boost ecotouristic value, and maintain water supplies.

In one such location, the Trifinio Region, an important transboundary region at the join of the borders of Guatemala, Honduras, and El Salvador, while considerable effort has been expended on the protection of coffee cropping, there is still limited knowledge of the extent, location, and importance of coffee land use. At its conception, the Trifinio Plan had been approved by its three participating nations (in 1987 in the aftermath of Central American civil wars) for the purposes of common forest and watershed protection, and eradication of poverty (Artiga, 2003). The region had been especially impacted by soil erosion due to high-elevation deforestation, due to planting of coffee on steep slopes. The initial Plan protected coffee crops and high-elevation biodiverse species. The follow-on “Trifinio 2” Plan reforested 4,500 ha to reduce the potential for flooding and spearheaded creation of a tri-national environmental mindset to protect the three main regional watersheds (Motagua, Ulua, and Lempa). Together the Plans envision sustainable coexistence, ecotourism, considerable protected forest and marine areas, as well as protected watersheds for millions downstream

which included mapping of water resources and monitoring of agricultural wastewater and runoff from coffee-growing regions with European assistance.

In one-such activity, a German Society for International Cooperation (GIZ)-funded project (ICP, 2015) “Tropical forest protection and watershed management in the Trifinio Region” worked directly with nearly 7000 coffee farming families, developing knowledge about how to adapt to regional water conservation needs to to permit increased aquifer recharge in sub-watersheds, as well as to protect and expand coffee agroforestry without compromising objectives. The forests and water program efforts also focused on improving understanding of the roles of women on the family farm, who not only rear and care for kids, but take care of the house, farm, and tend their cash (coffee) crops and subsistence production.

The root of the challenge (for the rural farmer in Trifinio Region to produce more efficiently, yet restrict activities to protect water supplies) is in its linkage to higher value agroforestry crops (particularly coffee) that optimally grow at specific elevation ranges, often overlapping with key micro-watershed recharge areas. Because mapping of coffee growing areas did not exist, educating coffee farmers required the creation of vinyl high-resolution Ikonos image maps that could be shared with farmers in their fields to explain the water resource protection situation at a scale more easily understood (Ingrid Hausinger, personal communication).

The only recent maps of coffee growing in the Trifinio Region stem from the national forest classifications of Rapid Eye 5-meter high-resolution images. The Ministries of Environment and Natural Resources (MARN) of Honduras, Guatemala, and El Salvador have each used these data to create recent mapped assessments of land cover and land use change (LCLUC). These maps include forest and crop types, some including shade coffee, suitable for developing Reference Emissions Levels (REL) that comply with national reporting requirements of signatories to the United Nations Framework Convention on Climate Change (UNFCCC) Paris Agreement.

MARN/El Salvador contracted a team to produce a REL map using these Rapid eye data (Center for Tropical Agricultural Research and Higher Education (CATIE), 2017). Making the map involved the mosaicking and classification of 11 overlapping least-cloud mosaics of Rapid Eye data to produce a pan-El Salvador forest cover map of 2015-2016, part

of which includes the Trifinio Region; a similar map was produced by using 54 classified Rapid Eye images from 2011.

The Trifinio Plans never envisioned a regional monitoring and mapping capability, though regional LCLUC maps (that include coffee as a land use class) were produced (CATHALAC, 2011) for natural resource planning and management. Through a framework agreement between the Trinational Territorial Information Systems Network (RITT), CTPT, local communities, university, and nonprofit agencies, those forest cover maps and additional data on agricultural and agroforestry stakeholders and natural resources management indicators have recently become available for download via the Trifinio Region information system and map server (Sistema de Información Territorial Trinacional or SINTET).

Coffee's role in the local economy

Coffee has been a feature of the Central American economy and farming families nearly three hundred years. Coffee (originally from Ethiopia) arrived in Central America almost 300 years ago (McCook, 2017). It has been used as currency. National governments have used it as an incentive for aid. In Honduras, the government linked it to county road-building money (Ian Cherrett personal communication). It is a choice crop of poorer farmers to be able to put more than corn on the table, to gain cash to expand land holdings (Fischer and Victor, 2014).

With globalization and increasing demand for coffee, farmers in Brazil and Vietnam increased production so well that in 1999-2002 supply exceeded demand causing the price to plummet causing the widely known "Coffee Crisis" that occurred across Central America and many farmers lost employment and went bankrupt. Since then, international prices have stabilized somewhat but started again to slip in 2010 and haven't yet recovered (CEDICAFE/ANACAFE, 2017). Locally during the period since the crisis, more than 50,000 new growers were added in Guatemala alone (Fischer and Victor, 2014). To complicate matters, the 2012 coffee rust (known locally as "roya") epidemic (devastated yields across the entire region) and is still impacting the area. In fact, today, coffee rust is impacting more than 31% of Trifinio Region's farms at a rate more than twice the national average (CEDICAFE/ANACAFE, 2017), while regionally, Honduras, El Salvador, and Nicaragua have extremely low rates of infestation. There are recent reports and commentary questioning

the sustainability of coffee in Guatemala with the rust, low prices, and high production costs (Bolanos, 2017). Development agencies working in the Trifinio region, on the other hand, report that coffee is the most important crop economically (Stiftung, 2013).

Recent statistics from Guatemala say that coffee production impacts 2.5 million people in Guatemala (CABI, 2017); that number could be higher in El Salvador, but the quantity of land involved in coffee production in the El Salvador portion of Trifinio Region is but a tiny fraction of that in Guatemala as El Salvador's land within the region is not optimally-located for coffee. In Honduras, there is an enormous amount of coffee in production now because the physical environment is optimal, there are government incentives for production, and though it is widely thought that the country has not suffered the recent coffee rust infestation as badly as its neighbors because they planted rust-resistant varieties, there are other opinions (Avelino, 2006).

Experience Researching Coffee

In Trifinio, two types of coffee are popular, sun and shade coffee (varieties of *Coffea arabica*); while these can be found at various elevations, coffees in shade are more likely found on the slopes of Trifinio's mountains near Copan, Honduras, for example, at elevations from 900-1300 meters (Smith, 2010). Sun coffee are most easily found on high-resolution satellite imagery when only 1-2 years of age, because the rows flow like pieces of textured ribbons. After two years of age and the bushes have grown together (visually) finding them is not so easily done with solely a single image, and Google Earth is a great tool to zoom in to find them (Figure 3.1).

Figure 3.1. Sun coffee (on a hill top near Esquipulas, Guatemala) is easily spotted because of their planting rows. Usually, they follow a dirt path which is never straight.

As noted in Figure 3.1, sun coffee cultivated by small producers is usually only visible from very close inspection with high-resolution imaging because the planting rows run helter-skelter and the color is spectrally confused with local forests (Cordero-Sancho and Sader, 2007; Ortega-Huerta et al., 2012; Gomez et al., 2010), though there are differences in textures. Shade coffee is not possible to map with coarse 30-meter Landsat 8 data, because it is grown underneath the forest canopy, and gaps in the canopy can only be seen with difficulty using very high-resolution data (e.g. two-meter Worldview 2 (LeLong et al, 2014) and/or Quickbird (Gomez et al, 2010) (Figure 3.2)). In Costa Rica, the shading component of coffee production is a native tree locally called “poro”. The branches of the poro are clipped before the dry season, permitting maximum sun and temperature for setting fruit. After three months, the branches grow back, and the dry season is over. In the Trifinio Region, farmers use native vegetation for shading, thus to a satellite imaging investigator it is almost possible to see the rows of coffee beneath the trees. Using Worldview 2 data, the rows of plantings do show in gaps and on the edges of the fields of trees. Within the canopy gaps, rows of planting and/or ploughing are visible (Figure 3.2).

Figure 3.2. Shade coffee can be seen in rows of through the canopy gaps.

Texture measurement tools have been used for almost 45 years to evaluate changes at the earth's surface for land cover mapping using satellite images to identify and measure production of and impacts to crops (Karakizi et al, 2016; Tang et al, 2015; Gomez et al, 2010), to assess urban zones (Zhang et al, , 2017), to estimate snowcover (Chen & E, 2007), to evaluate ice surfaces (Soh & Tsatsoulis, 1999), to assess forest structures (Ozdemir and Karnieli, 2011) and differences between arable and forest lands (Zhang et al, 2017; and Herold et al, 2004). In a study like this one, Chuang and Shiu (2016) analyzed Worldview 2 data with image pixels and object classification with significant field work using 21 different grey level co-occurrence matrix (GLCM) texture layers (Haralick, 1973) from Environment for Visualizing Images (ENVI) and classified them using machine learning techniques (random forest and support vector machine) to examine new possibilities for identifying and measuring very large industrial tea crops in fields of orderly rows in East Asia.

Though used to learn about images and also in raster format, GLCM measures are not images themselves (Hall-Beyer, 2007); they are raster results of convolution kernels that pass throughout an image. Applied to remotely sensed spectral data, the resulting GLCM layers help to predict the layout of landscape features (Hall-Beyer, 2017).

The sequences of texture differences found in landscapes of a remotely-sensed image range sometimes follow patterns that are naturally-derived or human-induced (for example, a field intermittently used for crops, burned, plowed, planted, and harvested may be characterized with texture measures). Under production, some fields are in recognizable rows, while more mature plants will have grown together, and very immature plants appear as a set of organized dots against a background of bare soil.

Using textures to evaluate coffee-growing in Trifinio with optical Google Earth images would only be possible when there are no clouds. In Trifinio, at high elevations, clouds are prevalent all year Thus, characterizing coffee with optical satellite images is only feasible on a clear day, and in the tropics, that is largely only possible in the dry season in Central America. Radio Detection and Ranging (RADAR) satellite images on the other hand can be used at any time of day or during most weather events. However, they present their own set of challenges for processing and understanding the signal. Historically, these images have been very expensive and difficult to process until recently.

The free proliferation of tools and images from the European Space Agency's (ESA) Sentinel constellation of satellites, and the distribution by ESA and the National Aeronautics and Space Administration (NASA) of Sentinel 1A and 1B, Synthetic Aperture Radars (SAR) and 2A (a Multispectral Imager with optical and infrared channels), is significantly helping to make remote imaging more accessible (ESA, 2010).

Several remotely-sensed investigations in coffee cropping regions, some in Central America, have tried to improve location of shade coffee, and assess regional coffee production and cropping challenges (Table 3.1).

Table 3.1. Some recent experience in coffee crop imaging in Central America

Classification Approach	Sensor	Classification Method	Authors
Unsupervised	Landsat 5/7; SPOT; MASTER & GeoEye-1	Unknown; ISODATA; Gaussian Mixture	Rueda et al., 2014; Cruz-Bello et al., 2011; Martignoni, 2011
Unsupervised	ALOS PALSAR1 & Landsat TM Combo	ISODATA	Montenegro and Atwood, 2010
Supervised	Landsat 5; SPOT; MASTER & GeoEye-1; Landsat ETM+	Maximum Likelihood (ML); ML; ML, Support Vector Machine (SVM); ML	Schmitt-Harsh, 2013; Martinez-Verduzco et al., 2012; Martignoni, 2011; Cordero-Sancho and Sader, 2007
Supervised	HyMap 2, Landsat 5/7	Random Forest, Decision Trees	Fagan et al., 2015
Supervised	Landsat 5 with thermal channel	Linear Spectral Mixture Analysis	Schmitt-Harsh et al., 2013

Accuracy of identification and classification, in these studies, was clearly an issue. None of the supervised classification studies surveyed surpassed overall accuracy of 86%. Cruz-Bello et al (2011) reported a range of accuracy of 94-97% using unsupervised classification with visual interpretation of panchromatic photos. To improve accuracy, Cordero-Sancho and Sader (2007) recommended using a subset of optical channels, combined with a vegetation index and a layer describing incident angle of solar energy. None of these studies used the new sensors, yet some of these have applied machine-learning classification algorithms.

Machine learning describes an applied set of instructions and improved results based on knowledge gains from repeated trials. The algorithms behind the instructions are not new. Randomforests (Breiman, 2001), has been around for two decades, for example; however, the development of these tools recently for “data mining” for the business, industrial, and thus image processing world is flourishing (Danilla, 2017) Applications languages, such as R (R Core Team, 2017) and Python (van Rossum, 2007) are advancing, as packages of instructions, such as “scikit-learn” (Pedregosa, 2011), “caret” (Kuhn et al, 2017), and others, learn and use other’s collective experiences for processing of remotely-sensed images.

Some remotely sensed studies have used Google Earth images to help validate remote places that experience severe cloud cover (Laso-Bayas et al, 2017). These images can be very useful especially in developing nations, despite reports of generally low accuracy among these, Google Maps, and Bing (See et al, 2017). They can be problematic, because one often hasn't very clear understanding of the date of acquisition.

Identification of remotely sensed features is carried out differently depending on image type and resolution. Each type is sensitive to different surfaces. Passive or optical sensors receive solar energy reflected off the surface of the Earth, whereas radar sensors that have large solar arrays to generate great power to actively send a signal to the Earth and collect billions of backscattered pulses of echoes. In images of 0.5-2m spatial resolution optical images, especially Worldview 2 data, many features can be readily identified by eye. On the other hand, some of the same features in 10-meter resolution Sentinel 2 and Landsat 8 images can be inferred by their shape, texture, and colors in red, green, and blue (RGB) channel presentations.

Active sensor backscatter in radar images, from nominally 10-meter resolution Sentinel 1A (Sentinel 1), SAR a C-band sensor are readily available for study of Trifinio Region's forest and agricultural lands. Other wavelengths of radar are available for the region, but at a high cost. The Japanese Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) 2, a L-band satellite-based sensor was launched in 2014. Its 23.4 cm wavelength, longer than Sentinel's 5.6 cm, is more useful to view tree structure deeper in the canopy than Sentinel. Sentinel can see objects as small as its wavelength, but is limited in terms of canopy penetration. The signal is responsive to textures, soil moisture, and look angle. High decibel backscatter is indicative of solid, manmade structures. Forest cover is usually bright as the signal bounces off all the branches in the upper canopy (known as 'volume scattering'), and relatively smooth water is dark because the signal bounces away and does not backscatter to the sensor. An interest of this study is that sun and shade will be detectable with C-band radar which is available almost every two weeks for free to be able to provide a monitoring capability not feasible previously.

Objectives

Data from Landsat 8, Sentinel 1, Sentinel 2, and Worldview 2 have only become available in the last few years, and due to their advanced specifications, it is important to test and analyze their use for neotropical monitoring. For example, these data may afford new opportunities for monitoring sun and shade coffee. In addition, it may be possible to use common texture measures combined with machine learning classifiers help us identify shade and sun coffees using these data. Specifically, my goals are three-fold, 1) to explore the differences between three relatively new sensors to identify coffee; 2) to test whether images of small producer sun, shade, adult and immature coffee production activities will exhibit differences in textures that can be classified and mapped; and 3) to compare the accuracy of five machine-learning classifiers to identify coffee with Landsat 8, Sentinel 1, Sentinel 2, and Worldview 2 satellite image data.

Methods

Study Region

The study takes place in the Trifinio Region, a 7400 km^2 transboundary region at the intersection of the borders of Guatemala, Honduras, and El Salvador. The local population of more than 600,000 is mostly rural, though the region has several large towns and cities. Agriculture is very important in the region; most land cover is impacted by farming and livestock production. The region has numerous protected areas (PAs) of various types and stages of growth (Munoz et al, in prep). One of these PAs, the Montecristo Trifinio Park, an important cloud forest of Central America, spans the tri-national borders of the participating nations. Extending outward from the PA boundaries, the rate of deforestation has varied due to coffee agroforestry development that competes with forest and biodiversity conservation (Schlesinger et al, 2017).

Data Availability and Acquisition

Trifinio experiences extended periods of cloud cover; this has prevented most viewing of the region by optical satellite imagery. There are only a few mostly optical multispectral Landsat satellite images of suitable resolution that are available each year with diminished cloud cover and only for the dry season (January-March). In most years, it takes an effort with

several images to mosaic together enough pixels to see Trifinio. A subset of the Trifinio Region was identified to carry out this study. The subset is delineated by the following World Geodetic Survey 1984 Datum coordinates:

Upper left Corner Longitude = -89.445445
Upper left Corner Latitude = 14.683407
Lower right Corner Longitude = 14.256615
Lower right Corner Latitude = -89.445445

Sentinel 1 synthetic aperture radar (SAR) images from the European Space Agency have been available for the Trifinio region in a limited fashion since mid-2014. SAR data can be used to collect data regardless of cloud cover. These data currently provide free 10m images every two weeks. Access to the Sentinel data is open; tutorials, software, training materials, and a user forum is free for all users (ESA, 2015).

SAR data differ from optical data in that their signal is actively transmitted and received in differing polarizations, horizontal, vertical, or a combination of the two; the optical signal is only passively received by a sensor passing by at the time of reflectance. Because the signal is not dependent on solar energy, it can be collected regardless of weather conditions or time of day; it is thus well suited for tropical research. The limitations of SAR include low positional accuracy. The Sentinel 1A data used in the current study is a C-band sensor; thus, its radar penetrates the vegetated canopy less than Advanced Land Observing Satellite (ALOS) PALSAR Phased Array type L-band Synthetic Aperture Radar (PALSAR) ’s L-band sensor. As a result, the reflected echoes interact more with the upper reaches of the canopy than the main stem of a tree. Because this C-band sensor has lower frequency, it has markedly reduced spatial accuracy (10m as opposed to a maximum of 3m of ALOS 2). But the Sentinel 1 data, tools, and methods for its processing are freely accessible, and available almost everywhere about every 2 weeks in the neotropics.

In this study, we used a single image from Landsat 8 and Sentinel 2, two adjacent images (in space and time) from Sentinel 1A, and eight adjacent images of Worldview 2 (Table 3.3) that make up the study subset (Figure 3.3).

Table 3.2. Image Dates from Sentinel 1 and 2, and Landsat 8 were acquired as near to each other in time possible.

Satellite	Date	Scene Identification
Landsat 8	4/9/2016	LC80190502016100LGN00
Sentinel 1A	4/6/2016	S1A_IW_GRDH_1SSV_20160406T114536_ 20160406T114601_010700_00FF50_5741.SAFE
Sentinel 1A	4/6/2016	S1A_IW_GRDH_1SSV_20160406T114601_ 20160406T114618_010700_00FF50_8D23.SAFE
Sentinel 2A	4/10/2016	S2A_OPER_PRD_MSIL1C_PDMC_20160412T201139 _R040_V20160410T163251_20160410T163251

Table 3.3. Eight Worldview 2 images were used in the study.

Date	Worldview 2 Identifier	Resolution
04/13/14	14apr13164944-m2as-500111330010_06_p005	2 meters
04/13/14	14apr13164945-m2as-500111330010_06_p006	2 meters
04/13/14	14apr13164923-m2as-500111330010_06_p013	2 meters
04/13/14	14apr13164924-m2as-500111330010_06_p014	2 meters
04/13/14	14apr13164926-m2as-500111330010_06_p015	2 meters
04/13/14	14apr13164927-m2as-500111330010_06_p016	2 meters
04/13/14	14apr13164942-m2as-500111330010_06_p003	2 meters
04/13/14	14apr13164943-m2as-500111330010_06_p004	2 meters

Acquisition of Landsat 8 occurred via the USGS Earth Explorer web portal, while download of Sentinel 1 data was facilitated by the Alaska Satellite Facility (NASA ASF). Sentinel 2 data were acquired and downloaded from European Space Agency (ESA). A full set of Worldview 2 data were available for the Trifinio Region through a partnership with US Fish & Wildlife. There is a two-year gap between the high and low-resolution images, therefore, the high-resolution images were used for identification of classes and training site facilitation.

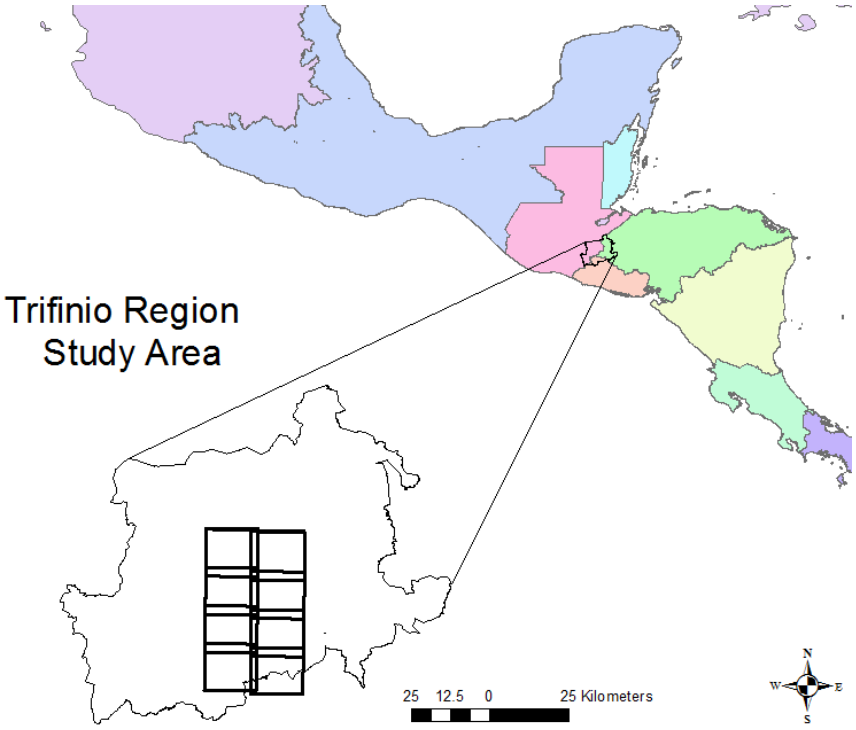


Figure 3.3 Trifinio Region Study Area (Central America)

Data Preprocessing

Satellite Data Preprocessing

Preprocessing of the optical data in units of proportional surface reflectance was accommodated by Clark Labs' TerrSet (Eastman, 1989); Landsat Climate Data Record imaging was accessed for Landsat 8 surface reflectance data. Sentinel Application Platform (SNAP) (*ESA Sentinel Application Platform v2.0.2*, 2017) was used to process radar images. Othocorrection of optical imagery was handled by the image vendors. The Worldview 2 data are orthocorrected with the Circular Error 90 Digital Elevation Model (CE90 DEM), Landsat data were orthorectified with ground control points to a level of Level 1 Terrain (L1T) correction. The L1T level of processing corrects pixels for the effect of terrain displacement caused by the sensing of the relief before passing by it. SNAP was used to correct Sentinel 1 data using a 3-arc second automatically downloaded chunk of a Shuttle Radar Topography Mission (SRTM) digital elevation model through a SNAP process called Range Doppler Terrain Correction. Range Doppler Terrain Correction orthocorrects each image pixel to the closest SRTM pixel, at the same time as correcting for the impact of the Doppler effect in the range direction (the direction of the satellite).

The Sentinel 1 SAR data used were of the Ground Range Detected (GRD) variety. These GRD data arrive preprocessed by ESA to remove the complex ("phase") data associated with each pixel, and they require fewer steps, but they have less data associated with them are less useful for certain applications. We applied precise orbital data, radiometric calibration, and speckle-filtering, followed by range doppler terrain correction. Application of precise orbital data permits correction of orbital offsets and degradation. We calibrated the image from raw pulses to units of backscatter in sigma0, sometimes called "sigma nought". Once calibrated, an image can be compared with other calibrated images regardless of incidence angles or brightness levels.

Speckle filtering reduces the impacts of the constructive and destructive interference inherent in radar images. These 'speckle' pulses result from the effect of many backscattering elements. The effect or impact of speckle within an image can either be reduced or enhanced by passing processing windows of known dimensions over the entire image and performing the same filter algorithm. This can occur across all pixels, or adaptively to mostly carry out one task on pixels of similar slope, for example.

SAR Filter Evaluation

Ozdarici and Akeyurek (2010) recommended an evaluation of differences between means and standard deviations in pre-filtered and post-filtered SAR data to determine the best type and dimensions of filter to reduce image speckle. The objective with speckle filtering is to generally reduce the backscatter noise while maintaining a high correlation between the original and filtered SAR image. These two statistics are used to derive two measures called Quality and Equivalent Number of Looks (ENL). If edges are a necessary component of the landscape, adaptive filters are designed to reduce speckle but preserve edges. Previous work indicates how to determine the optimal size of filter to use in processing various datasets (Ozdarici and Akeyurek (2010); Anys & He, 1995). We tried 3x3, 5x5, 7x7, and 9x9 sized filters of 10 different varieties to evaluate the optimal filter type and size for delineating coffee plantations from other land use and land cover types (Table A.2), and ultimately chose a Lee 3x3 filter.

Topographic Effects

Terrain flattening is recommended for applications involving SAR image classification. Flattening can eliminate the impact of topography on SAR backscatter. SAR backscatter variation is manifested by either of two SAR geometric viewing-related issues: foreshortening (when the angle of incidence is not 90 degrees) and layover (when the top of a slope is nearer to the sensor platform than the bottom and the signal returns more quickly). Range Doppler Terrain Correction was performed in the Sentinel Application (SNAP) software.

Clouds and Shadows

We masked clouds and cloud shadows in the Landsat 8 and WorldView data prior to image analysis. In Landsat images, we used the Fmask layer (Zhu & Woodcock, 2012) to mask out the clouds.

Image Segmentation

Traditionally practiced, remotely-sensed image analysis attempts to categorize landscape feature knowledge and classify images using supervised or unsupervised grouping of similar pixel values in spectral space, separated using knowledge categories. Geospatial object-based image analysis similarly uses ‘rule sets’ to separate similar pixels by their knowledge categories of shape, length, pattern, texture, and spectral qualities (Whiteside & Bartolo, 2014; Blaschke, 2010) (into vector objects or segments – a process known as “segmentation”). All tiles of the Worldview 2 images (Table 3.3) were segmented into vector objects, called ‘segments’ using TerrSet v. 18.31 to ultimately create raster and vector shapefiles that could be used as training sites for both object-driven and image-based classification methods. TerrSet’s segmentation method uses a watershed approach. We used a 3x3 window with a similarity tolerance of 50, and equal weights for the full set of input bands, mean, and variance factors, capturing 20 classes of objects (Table A.1).

Data Analysis

Mapping Vegetation

Four main types of coffee exist in the Trifinio Region, including immature coffee with low canopy cover, mature coffee (bushes are darker and taller than others and also in rows, but more dispersed than immature plants, ‘sun coffee’ of various ages grown in full sun, and ‘shade coffee’ grown next to or under other tree species. Farmers often chop the upper branches of the shade components of coffee growing in shadows in the beginning of the dry season and then this growth returns with the onset of rains at the start of the rainy period.

Some kinds of trees in high-resolution images are more obviously identified than others. The crowns of certain species of coniferous trees are pointy and easier to detect than some, and their degree of reflectance is less than broadleaved varieties in the NIR band, but in the dry season, when some leaves are off, and others are on, we decided not to try to determine what was a tree. All likely forested objects not identified as shrubs, sun or shade coffee bushes, or large tree crowns were assigned to the forest class. In this study, large single tree crowns were often identified as the only object of a segment. These usually occurred at the edges of coffee fields, as these large canopy trees are mostly outside of a field or in a pasture. All single tree crowns were reclassified to the forest class. Shrubs are common, but individual shrubs can only be identified on high- resolution images, probably because they only begin to become identifiable at those scales. The shrub class exists on the El Salvador forest cover map, but not on the maps of Guatemala nor Honduras. Shrubs are identifiable by their stature, dispersed locations (not in rows), often at the edge of or near the edge of fields or fallow. One of the problems with the class is that height (which sometimes determines the difference between a shrub and a tree) is not easily distinguishable in passive imagery. We determined shrubs from forest by lack of associated shadow, as trees invariably have shadows. Aside from the various coffee classifications, forest, and shrub classes, all other land use and land cover types were unclassified.

Within the subset area chosen for analysis, differing classes and quantities of landcover classes were found depending on the resolution of the data used to evaluate. Image segmentation of the Worldview 2 data permitted classification of 20 different classes (Table A.1).

Of the 20 landcover classes mapped across the images used in the study, only five classes (shade, sun, adult, and immature coffees, as well as forest/bush cover) were critical to the first part of this work, thus all statistical analysis for the first question were related to determining differences in the perception of coffee by the new sensors. These five classes were compared using logistic regression in R to the seven TerrSet texture types and four grey-level co-occurrence matrix measures involving four different satellite sensors (Table 3.2).

Training Sites & Validation Maps

We generated training sites to conduct supervised classification of the imagery. Some training sites were identified in existing maps acquired from the three participating countries of the Trifinio Region (REDD/CCAD-GIZ, 2014; Jimenez, 2014; and Catalan et al, 2015). Each of these maps were independently created using five-meter Rapid Eye optical satellite data. The completed legend has the classes of: Pasture, Coffee, Forest, Agriculture, Bushes, Fruit, Urban, Water, Unclassified, and NoData.

Validation of the study region map was carried out using Google Earth Pro (version 7.3.0.3832, Google, 2017) with historical imagery data activated. Three polygons of specifically coffee agroforestry sites were randomly selected to be about 150 hectares each, and located in one of the countries. A random sample set of 100 locations were identified for each polygon, and each sample was identified visually as to whether it was shade, sun, or not coffee (an average of 68% of the combined national map polygons were identified as either shade or sun coffee). Additional to the Google Earth mapping activity, locational data in geographic coordinates were available for about 7,000 coffee and agroforestry cooperative members and their farm locations were downloaded from the Trifinio Region SINTET Geoserver in July 2017. These agroforestry participants were mapped and concentrated within the Trifinio Region's 15 key sub-watersheds. A subset of these agroforestry farm locating polygons (87) overlapped with our study region.

We created a map of sun and shade coffees to be able to compare the usefulness of the optical and radar images. We used eight scenes of WV2 data at 2m resolution in the middle of the Trifinio Region (boxes in center of Figure 3.3) centered on a majority number of coffee polygons identified by the combined 5m resolution mapping done by the three country environmental agencies. The same areas had been mapped with Landsat 8 at 30m and by

Sentinel 1A SAR and Sentinel 2A Multi-Spectral optical data at 10m-resolution each. A tile structure was created to accommodate the combined images. While the 10m medium resolution data could use the entire area, the high-resolution data needed to be mapped separately because collectively these were too large for any of the software and hardware processing tools available. The sun and shade class polygons mapped from Worldview 2 were combined in a single map for the study area and these points were validated using random samples and the Google Earth procedure described above.

Vegetation Indices & Others

We calculated the normalized difference vegetation index (NDVI; Rouse et al., 1973) to characterize differences between pairs of image bands in remote sensing image processing. Cordero-Sancho and Sader (2007) recommended that adding NDVI for classification of coffee was useful to improve classification accuracy (Equation 1) when combined with a subset of Landsat Thematic Mapper (TM) bands and a layer characterizing the effective incidence angle (EIA) of reflected solar energy (Holben and Justice, 1980). In this work we chose Landsat 8 bands comparable to those used on Landsat 5 in the original research (Barsi et al. 2014). The incidence angle is the difference between the solar declination angle and the solar normal, and can be calculated from a combination of trigonometry and a digital elevation model, such as Shuttle RADAR Topography Mission (SRTM) data (Farr et al, 2007). The new Angle Coefficient file that comes with the Landsat Collection 1 Landsat data from USGS contain angular data to calculate solar viewing angles on a pixel by pixel basis as opposed to only using the scene center Sun Elevation and Sun Azimuth reported in the Metadata Library (MTL) file, but the equations and tools for the new Solar and View Angle Generation Algorithm have not yet been made available (USGS, 2017).

$$\cos(i) = \cos(S\_z)\ \cos(\beta) + \sin(S\_z)(\sin(\beta)\cos(a\_z - aspect)) \qquad \text{(Equation 3.1)}$$

where

S_z = solar zenith angle (^{o})

β = slope (^{o})

A_z = solar azimuth angle (^{o})

(Cordero-Sancho and Sader (2007)

Slope and Aspect were calculated in degrees. The Solar Zenith Angle is equivalent to the difference between 90 degrees and the Solar Elevation angle in degrees. The Solar Azimuth Angle (also in degrees) is equivalent to the SUN_AZIMUTH parameter in the Landsat MTL file.

Textures

Texture measures used in this study were found in standard image processing software (e.g. ENVI and TerrSet). Texture measures are comprised of four groups, geometric, statistical, model-based, and signal-related (Ojala and Pietikainen, 2002). Geometric and statistical measures are commonly used. Hall-Beyer (2017) reported in her GLCM Tutorial the importance of an assessing band to band correlation by image type to ascertain which texture measures are most independent and not correlated with another to be able to reduce the quantity of bands, amount of work, and increase the value of classification outputs.

A total of 15 different types of texture measures (seven texture patterns from TerrSet and eight GLCMs from ENVI) were compared for each band of every input image. These included: Relative Richness, Fragmentation Index, Diversity, Dominance Index, Number of different classes (NDC), Binary comparison matrix (BCM), and Center versus Neighbors (CVN) (consult TerrSet for specific information on the textural variability measure algorithms) (Eastman, 1989); as well as the GLCM measures mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation (Harelick, 1973). A 3x3 window was applied to reduce resolution the least, and consider the size of the coffee plant.

Visualizations supporting exploratory data analyses (using qqplots, boxplots, and histograms), followed by omnibus normality measures (e.g. Lilliefors, Kolmogorov-Smirnov (p-value < 2.2e-16)) from the R Analysis of Over-Dispersed Data “aod” package (version 1.3) (Lesnoff and Lancelot, 2012) that supported a conclusion that these texture data and mapped forest and coffee classifications were not from a normal distribution. Attempts to transform these data and reassess normality were undertaken with a randomly stratified sample of nearly 73,000 forested land cover and coffee land use class values, which were subjected to log, square root, and cubic transformations in rigorous in R. The cubic transformation was found to impact histogram structure considerably.

A literature review of GLCM texture measures (Hall-Beyer, 2017) recommended the inclusion of mean and contrast measures, plus the addition of two or three others and no more than that to avoid confusion. An examination of these measures showed that entropy and correlation could be useful to pursue. Therefore, these four measures were included in subsequent logistic regressions in addition to a set of seven pattern measures from TerrSet to assess significant coffee relationships. General Linear Model logistic regressions in R were used to assess potential texture relationships with forest and coffee, and several significant relationships were noted (Table 3.2), but were invalidated in subsequent testing (with Hosmer.test and Wald.test, deviance measure assessment, and relative operating characteristic (ROC) estimation in TerrSet).

Data Reduction

Including the 15 texture layers for each of the 30 bands of spectral and radar backscatter data, there were a total of 450 prospective raster predictor variables for use in the study's land cover and land use classifications. Additionally, there are 7 different window sizes, 5 different offset levels and 3 offset distances that need to be set when preparing any texture layer using TerrSet or ENVI, for a grand total of 47,250 potential classifications to carry out our study. Thus, rigorous band selection was carried out using the "cor.test" package in R (v.3.3.3) to create correlation matrices and to determine collinearity (Tables A.3, A.4, A.5, and A.11). Principal components analysis with Worldview2, Sentinel 2, and Landsat 8 images in TerrSet reduce inputs to the least quantity of principal components to make the research even more efficient (Table A.6); two principal components of Worldview were used in each tile to produce training site polygons that were then combined later to develop classifications. In the end, with the selection of a single GLCM processing window dimension, we had reduced our requirements to 45 different model runs. A literature review of texture measures and their applications (Hall-Beyer, 2017; Kailath, 1967) suggested that not all texture measures would be useful as some are more correlated with different spectral bands than others. Subsequently all 15 original texture measures were regressed (using logistic regression in R) against nearly 73,000 stratified random samples of binary classifications of the four Worldview 2 coffee and forest classes to seek only those texture

measures showing significant relationships (Table 3.1). Four pattern measures from TerrSet and four GLCM measures from ENVI were selected for further research (Hall-Beyer, 2017).

Machine Learning Classifier Selection

Six classifiers were selected to model our training sites (Neural Network (nnet), Support Vector Machine (svm), Random Forest (rf), Maximum Likelihood (mlc), Naïve Bayes (n_b), and k-Nearest Neighbor (knn)). We selected these classifiers because they each accommodate non-parametric data. We used the R programming environment (version 3.3.3, 2017, "Another Canoe", R Core Team 2017) for implementing the classifiers. All runs used the same set of input parameters, though the classifier mnemonic and input and output files needed to be changed for each run. The superclass function is from the RStoolbox (Leutner and Horning, 2017), and uses 'train' from the caret package (Kuhn et al, 2017). The code is parallelized to take advantage of multiple cores. The input files for model runs used the TerrSet IDRISI file structure for input and output. These codes can be easily changed to accommodate other file systems and structures using existing infrastructure or additional R packages (such as the Geospatial Data Abstraction Language package for R, called "Rgdal" (Bivand et al, 2017)).

Results

The key results of this study are that texture variables show statistically significant association with shade coffee and forest cover, and that machine learning classifiers can produce coffee mapping results that are comparable to those produced in previous research efforts.

Logistic regression modeling

The logistic regression modeling of the coffee and forest type classes in R (version 3.3.3, R Core Team 2017) initially showed significant relationships with three of TerrSet's texture measures (including Relative-Richness, Diversity, and Dominance Index (Turner, M.G., 1989). However, follow-on assessment with strength tests (Hosmer-Lemeshow from the 'ResourceSelection' package in R (Lele, et al., 2017) and Wald test from the 'aod' package in R (Lesnoff & Lancelot, 2012)), invalidated almost all of the significant

relationships, except for three models (Table 3.4) that demonstrated significance at the 0.90 and 0.99 levels for shade coffee and forest cover (see R code in Table A.8). The deviance model of logistic regression tests the null hypothesis that the deviance measures of the constant model are the same as the deviance measures of the residual model. In this case, the p-values of the logistic regression are significant when the null hypothesis for deviance is rejected. In this case (Table 3.4), the null hypothesis was rejected in three models for shade coffee and forest cover by relative richness and diversity texture measures.

Table 3.4. Logistic regression results of predicting coffee targets with independent texture layers based on Sentinel 1 & Deviance P-Values

Target	Predictor Texture	P-Value, Signif. Level	Confidence Interval	Deviance P-Value
Shade Coffee	Relative Richness	0.00901, **	[.0024, .464]	0.006737 **
	Diversity	0.00149, **	[-0.716,-0.0719]	
Forest	Diversity	.0799, .	[0.14, 0.41]	1.30E-10
Forest	Relative Richness	0.0936, .	[0023-.2732];	2.29E-11
Forest	Diversity	0.0772, .	[0157-.4099]	

*Significance level: p<0.05**; p<0.1”.”*

Subsequent logistic regression modeling of the vegetative type classes in R (version 3.3.3, R Core Team 2017) as binomial targets of 0 or 1 as a function of the seven texture layers from Terrset showed significant relationships between the coffee types identified and four landscape ecology variability pattern measures (including Relative-Richness, Diversity, and Dominance Index (from Turner, M.G., 1989), and Fragmentation Index (Monmonier, M.S., 1974) created by assessing sensor bands with the PATTERN.EXE program in Terrset, where the equations are explained).

Logistic Regression Results of Texture Analyses.

Logistic regression was sought to find relationships between transformed predictors texture layers and the target features in R. Potential relationships were invalidated using Hosmer-Lemeshow from the ‘ResourceSelection’ package (Lele, et al., 2017), the Wald test from the ‘aod’ package (Lesnoff & Lancelot, 2012), as well as the Relative Operating

Characteristic (ROC), as well as the calculation of confidence intervals for the significant relationships.

Kappa Commission/Omission

Validation of all machine learning classifications compared shade and sun coffee classifications with the validation map (Table 3.5). At this point, however, shade coffee could only be compared between Landsat 8 and Sentinel 2, because there were no significant relationships found between the SAR texture measures and coffee situations.

Table 3.5. Filter Type and Size Evaluation

MEAN	Reduce/	Kernel				MEAN DIFFERENCE (POST-FILTERING)				BEST SCORE
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.1244	0.1244	0.1244	0.1244	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
Median	R	0.1171	0.1103	0.1061	0.1033	-0.0073	-0.0141	-0.0183	-0.0211	**(larger value**
Frost	D	0.1236	0.1228	0.1229	0.1232	-0.0008	-0.0016	-0.0015	-0.0012	**= higher**
Gamma Map	R	0.1243	0.1242	0.124	0.124	**-0.0001**	**-0.0002**	-0.0004	-0.0004	**quality)**
Lee	D	0.1243	0.1242	0.124	0.124	**-0.0001**	**-0.0002**	-0.0004	-0.0004	**Lee**
Lee Sigma	D	XXXX	0.1258	0.1245	0.1234	XXXX	**0.0014**	**0.0001**	**-0.0010**	
IDAN50	R	0.1087				-0.0157				
IDAN75	R	0.1067				-0.0177				
IDAN100	R	0.1055				-0.0189				
RefinedLee	R	0.1176				-0.0068				
SIGMA										
	Reduce/	Kernel				SIGMA DIFFERENCE (POST-FILTERING)				
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.144	0.116	0.0999	0.893	-0.0421	-0.0701	-0.0862	0.7069	
Median	R	0.1227	0.0866	0.0712	0.0622	-0.0634	**-0.0995**	**-0.1149**	-0.1239	**(smaller value**
Frost	D	0.1449	0.1573	0.1629	0.164	-0.0412	-0.0288	-0.0232	-0.0221	**= higher**
Gamma Map	R	0.1435	0.1141	0.1055	0.0943	-0.0426	-0.0720	-0.0806	**-0.0918**	**quality)**
Lee	D	0.1437	0.1158	0.1015	0.1038	-0.0424	-0.0703	-0.0846	-0.0823	**Lee**
Lee Sigma	D	XXXX	0.1769	0.5112	0.172	XXXX	-0.0092	**0.3251**	-0.0141	
IDAN50	R	0.089				**-0.0971**				
IDAN75	R	0.0852				**-0.1009**				
IDAN100	R	0.083				**-0.1031**				
RefinedLee	R	0.1168				-0.0693				
ENL										
	Reduce/	Kernel				ENL DIFFERENCE (POST-FILTERING)				
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.7459	1.1492	1.549	1.9375	0.2995	0.7028	**1.1026**	**1.4911**	
Median	R	0.9115	1.6215	2.2172	2.7564	0.4651	**1.1751**	**1.7708**	**2.31**	**(larger value**
Frost	D	0.7272	0.6095	0.5692	0.5642	0.2808	0.1631	0.1228	0.1178	**= higher**
Gamma Map	R	0.75	1.1819	1.3772	1.7128	0.3036	0.7355	0.9308	**1.2664**	**quality)**
Lee	D	0.7483	1.1502	1.4934	1.4273	0.3019	0.7038	**1.047**	**0.9809**	**Lee**
Lee Sigma	D	XXXX	0.5054	0.5112	0.5147	XXXX	0.059	0.0648	0.0683	
IDAN50	R	1.4922				**1.0458**				
IDAN75	R	1.568				**1.1216**				
IDAN100	R	1.6165				**1.1701**				
RefinedLee	R	1.0147				0.5683				

Classification accuracy of effective incidence angle (EIA) product

Thermal infrared sensor (TIRS) 2 from Landsat 8 combined with bands 4, 5, 6, 7, NDVI, and EIA, when compared with TIRS1 using a maximum likelihood classifier, offered the highest accuracy in the identification of coffee (Table 3.6).

Table 3.6. Image accuracies, errors of omission, errors of commission, and kappa index of agreement for the coffee class.

Image Source	Classifier	OverallA	ErrorO	ErrorC	Kappa Index by Class*
Landsat 8	knn	71%	21%	80%	31%
Landsat 8	mlc	87%	12%	90%	60%
Landsat 8	n_b	76%	15%	86%	21%
Landsat 8	nnet	73%	26%	88%	31%
Landsat 8	rf	71%	15%	84%	73%
Landsat 8	svm	75%	31%	86%	46%
Sentinel 2	knn	73%	35%	86%	43%
Sentinel 2	mlc	77%	42%	88%	31%
Sentinel 2	n_b	80%	37%	85%	43%
Sentinel 2	nnet	69%	46%	85%	35%
Sentinel 2	rf	76%	30%	86%	47%
Sentinel 2	svm	69%	46%	84%	35%

OverallA = Overall Accuracy

ErrorO = Error of Omission

ErrorC = Error of Comission

*Kappa index of agreement per classifier for the coffee class)

Discussion

Research-achieved accuracies

Drawing a line down the middle of the overall accuracies reported by previous research teams (Table 3.7) yields about 62% accuracy (eliminating the estimated data of Langford and Bell (1997) who applied the methods of Card (1982) that may not be compatible with other reported efforts); averaging the maximum overall accuracies reported by these research teams

yields about 82%. Our own average overall accuracy of 12 classifications of Sentinel 2 and Landsat 8 data (Table 3.6, with confusion matrices found in Table A.10) yielded about 75%, which puts us at a little better than the middle of the pack, even if we also add in the results of the hybrid classifications (Table 3.8) using NDVI, TIRS, and EIA in Landsat 8. Our errors of omission were smaller than previous work, but our errors of commission were of similar amounts. These results (Table 3.6) indicate that there is substantial room for improvement, especially on class identification.

Table 3.7. A review of accuracies and errors in coffee literature over 20 years

Research Team	Overall Accuracy (Percent)	Error of Comission/ User's	Error of Omission / Producer's	Data Type
Bolanos	72-75%	Na	na	Landsat
Cordero-Sancho and Sader	63 %	92%; 86%	55%; 68%	Landsat
Cruz-Bello et al	93%	Na	na	SPOT
Gomez et al	61-83%	Na	na	Quickbird
Langford and Bell	38-59%	*	*	Landsat
Martinez-Verduzco	73%	60-83%	92%	SPOT
Mukashema et al	83-97%	Na	na	Quickbird
Rahman and Sumantyo	83% (forest)	83%	87%	SIR-C, Alos Palsar
Schmitt-Harsh	73-86%	67%	89%	Landsat TM

Table 3.8. Classification accuracy achieved by inclusion of NDVI and EIA with each of the short-wave infrared (SWIR) bands.

Bands Used	Classification Accuracy Achieved
4,5,6, 7, TIRS1, NDVI, EIA	78%
4,5,6, 7, TIRS2, NDVI, EIA	79%

Classification accuracy achieved by inclusion of NDVI and EIA with each of the thermal-infrared sensor (TIRS) bands (Bands 10 and 11 in Landsat 8).

We might be able to improve in the way in which we view and use these indicators (e.g. errors of omission, errors of comission). We tend to use them statically to measure the quality of our work, however, new research (Lu et al, 2014) suggests that these values are only an intermediary output that can be put to work to determine achievable accuracy and overall levels for specific cover classes, suggesting that it isn't that we necessarily failed to produce, but that we failed to identify a maximum level of accuracy that can be achieved given a certain set number of inputs and which set of layers (spectral or textural) need to be put together to create an expected level of accuracy.

Sensor differences and added ancillary data for coffee classifications

The sensor resolution of Worldview 2 was degraded pre-classification (from 2x2 meters to that of Sentinel 2. Landsat 8 was expanded post-classification (from 30 to 10 meters) to accommodate their comparisons with the Sentinel 1, Sentinel 2 data sets and degraded Worldview 2. This almost certainly had negative consequences, especially for the experience of Worldview 2 which showed the worst performance. There was one run with a non-convergence of the maximum likelihood classifier. SAR surprisingly with its single polarization did not exhibit the worst performance for identification of coffee.

The study date was fixed at a time when co-polarized data had not yet been made available for Central America. It has been proven in other studies that multi-polarized and co-polarized SAR offers a much better classification performance when compared with use of any single polarization (VV in this study). In prior work (with ALOS PALSAR 1 data) in the study region, both horizontal send and horizontal-receive (HH) and horizontal send, vertical receive (HV) data were available (during 2006-2011), and the experience with HV co-polarization was more fruitful for the identification of forest cover than only using HH. Co-polarized (VV and VH) data are now available for the Trifinio Region. Landsat 8 data showed the most accurate results, followed by Sentinel 1.

The inclusion of the EIA with optical, infrared, and vegetation index layers demonstrated that TIRS2 improved the classification of coffee, using the maximum likelihood classifier; yet it offered no greater classification accuracy value than that experienced with the other sources and techniques. It seems that inclusion of the NDVI and EIA layers were an aid to classification (Cordero-Sancho and Sader, 2007), but more work is necessary there to

determine which part of those layers are the most important contribution to the higher accuracy experience. Holben and Justice (1980) offered the EIA layer as a correction for the effect of angle of a slope and the direction in which it faces, as sunny and bright east-facing slopes with different radiances are invariably classified differently than those showing up darker due to slopes facing other directions. The inclusion of the EIA with optical, shortwave infrared, and vegetation index layers t perform better than the top two classifiers (MLC and N_B) in terms of overall accuracy. Inclusion of the NDVI and EIA proved to be a marginal improvement for the classification of Landsat TM layers (Cordero-Sancho and Sader, 2007); they wrote that land cover classifications that included both NDVI and EIA only improved when topographic data were not included as part of the classification. Sesnie et al, (2008) tested four vegetation indices (VI) to improve separability of woodland classes being evaluated by machine learning classifiers support vector machines and random forests classifiers to separate forest types in tropical Costa Rica. Sesnie et al (2008) wrote that one of the four VIs increased spectral separability more than another index, and two of the four spectrally separated reforestation layers from regrowth layers, while the remaining VI type (Enhanced Vegetation Index) had poor performance. Sesnie et al's team also found that applying an elevation layer to the classification achieved greater class separability, Elevation and slope layers (from the Shuttle Radar Topography Mission – SRTMGL1) were included in this study's logistic regression testing to develop texture predictors, but though p-values achieved were initially found significant, deviance testing invalidated any potential relationships. More work is necessary on this topic to determine which part(s) of these layers contributes more to classification accuracy.

A central theme that emerges is the choice of predictors and predictor bands. We reduced the band set using band to band correlations and principal components among some methods to direct our choices. Correlation among optical bands was high among the medium resolution sensors (Sentinel 2 (Table 3.9), Landsat 8 (Table 3.9), and Worldview 2 (Table 3.11) with considerable redundancy seen. This is certainly important for cost and data reduction (time and money) and improving classifications. The GLCM correlation by Landsat 8 bands was interesting (Figure A.3), as the infrared bands showed very low correlation. Not enough classifications were undertaken to know whether this was useful or not.

Table 3.9. Sentinel 2A Band Correlation Matrix

	B2	B2	B4	B8
B2	1	**0.97**	**0.96**	0.1
B3	**0.97**	1	**0.95**	0.24
B4	**0.96**	**0.95**	1	0.07
B8	0.1	0.24	0.07	1

Table 3.10. Landsat 8 Band Correlation Matrix

	B1	B2	B3	B4	B5	B6	B7	B9
B1	1	**0.99**	**0.96**	**0.95**	0.12	**0.82**	**0.87**	-0.2
B2	0.99	1	**0.97**	**0.97**	0.13	**0.84**	**0.9**	-0.17
B3	0.96	0.97	1	**0.97**	0.27	**0.85**	**0.88**	-0.12
B4	0.95	0.97	0.97	1	0.13	**0.9**	**0.93**	-0.11
B5	0.12	0.13	0.27	0.13	1	0.24	0.1	0.04
B6	0.82	0.84	0.85	0.9	0.24	1	**0.96**	-0.11
B7	0.87	0.9	0.88	0.93	0.1	0.96	1	-0.09
B9	-0.2	-0.17	-0.12	-0.11	0.04	-0.11	-0.09	1

Table 3.11. Worldview 2 Correlation Frequency Matrix (by band 1-8).

	B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**
B1		8	8	8	2	0	0	0
B2			8	8	5	0	0	0
B3				8	8	2	0	0
B4					8	0	0	0
B5						0	0	0
B6							8	8
B7								8

Logistic regression strength

There are several areas of concern with this study. The study was premised on the idea that GLCM and pattern textures could be used in lieu of spectral variables that are confused because of closeness of coffee leaf color with that of surrounding native vegetation. It was supposed that the four types of coffee structure (sun, shade, adult, immature) variables related to how the plant is grown, its physical position on the landscape and its age would be able to define changes to aid in classification. The logistic regressions demonstrated that there were

initially significant findings for three variability texture measures (relative richness, diversity, and dominance index), but a closer comparison of the differences between the deviance values in the residual model versus the null model invalidated the calculated p-values, and subsequent checks of those p-values with additional measures (Hosmer Lemeshow and Wald) gave more evidence for that conclusion. At the end of the R investigation, two quick additional logistic regressions (from the LOGISTICREG.EXE program with Terrset) showed very low relative operating characteristic (ROC) scores indicating, as found previously, that the outcome of the relationship between the texture predictors and the coffee variable targets was no more than a chance occurrence.

Texture and coffee

It was assumed from the start because of successes reported in the literature (Zhou et al, 2017; Herold et al, 2004) with row crops in China (e.g. wheat, corn) and in Kenya (with tea plantations) that texture measures created for coffee agroforests were going to be easily measured and seen to be related to different statures of coffee bushes. It was also assumed at the outset that texture measures from optical and radar bands could be used to aid in classification of coffee as there is a wide variety of literature supporting the topic. However, almost no studies reported on a relationship between shade and sun coffee and measured texture, and certainly not by using SAR. Two studies noted they had used GLCM in their work on coffee, but did not share any result (Rueda et al, 2014). Herold et al. (2004) reported on the use of GLCM and SAR in locations other than commercial agriculture, but not with coffee. A study most like this one, took place on the Pacific Island of Caledonia, and used 48 GLCM measures in filtered neighborhood windows of 99x99 pixels with optical sensors and machine classifiers to develop relationships between canopy heterogeneity and tree crowns to be able to predict locations of shade coffee (Gomez et al, 2010). Most of the coffee agroforests in the Trifinio Region are located on variable landscapes; some are very inaccessible, located on slopes or on hilltops.

There may yet be a relationship between texture variables and coffee, despite one not found in this examination, for a variety of reasons (e.g. texture window dimensions and overall reduction of texture because of the timing of the speckle reduction, as well as insufficient ground data for classification and verification due partially to a lack an onsite

component and insufficient investigator site knowledge). To be more specific, there is little consensus for setting the thresholds. Hall-Beyer (2017) noted that a simple assessment of the best filter type, convolution kernel size, and considered parameters could present a user with as many as 6300 different filter options to carry out a study to follow to determine which filter to use at what dimensions; nonetheless, a literature review found that some researchers stress comparisons of an image quality measure (based on differences measured in pre-filtering and post-filtering image means and standard deviations), with the results of a commonly used equation in radar image processing, called the Equivalent Number of Looks (ENL) (Ozdarici and Akeyurek, 2010). The idea is that the filter-based speckle reduction process should not overly impact the image mean nor standard deviation. A comparison of 10 filters with 4 window dimensions and 3 neighborhoods demonstrated that the Lee filter (Lee, 1983) used at 3x3 dimensions was the main dimensions chosen for this study (Table 3.6). Strengths of training site locations with coffee textures could have been addressed with Jeffries-Matushita.

This decision to use the Lee filter took into consideration the impact of the filter on the size of the object to be imaged (small and large coffee bushes) and the fact that the production is often small-footprint agroforestry, many times haphazardly planted (as opposed to large-scale industrial farming occurring in orderly rows. An adaptive filter that is edge preserving would not likely add value to images of small-scale agroforestry. Works from other studies conflict with this decision. Two investigations reported that the chosen kernel size impacts coefficient of variation (CV) impacting the ability of crops being classified (Tonye, 2002; Anys & He, 1995). Anys & He (1995) noted that the smaller (3x3) window caused the CV to be too broad, and largest filter they studied (13x13) caused the CV to be too small to distinguish differences among classes and that the 7x7 filter they chose was a compromise for that situation. In other studies, it was noted that the size of filter could be very large. A review of filters use showed some at 21x21 and still others at 99x99 (Brito, 2003).

Validation of Training Site Maps

The maps locating coffee agroforestry plantations were each derived from 5x5-meter Rapideye data for the three countries by separate agencies in different periods without knowledge of the others' efforts; the three products are independent; they were not intended to be used in conjunction with each other. Almost no other data were found to corroborate

these map classes, however, there are some possibilities. For example, the CATHALAC map (2010) was produced before the RapidEye images were even acquired. Additionally, nearly 7,000 members of coffee cooperatives and agroforestry associations were mapped in the last few years by the coffee cooperative associations of the three countries, working with Trifinio Commission projects, and the points associated with these mapping efforts were recently made available and downloaded from the CTPT's SINTET Geoserver. Within these datasets organized by watershed and association name are coffee farmer names, their coffee crop local name, the GPS locations of their fields for each of the sub-watersheds identified as important for aquifer recharge for water supply distributed by the three main watersheds (Motagua, Lempa, and Ulua rivers). The farm sites are distributed by geographically among the three countries by watershed, and the efforts to collect those data were independent of the three country forest mapping efforts that created the coffee map used as a training site input for image classifications in this study. The second part of the issue relates to lack of an onsite component, a limitation, because there was insufficient material to know with full confidence that the locations chosen as training sites for shade and sun and adult coffee were adequately selected (the immature coffee was very obvious visually seen on undegraded Worldview 2 data). The two methods used to determine these classes were visual inspection of high-resolution Worldview 2 imagery and the 1-meter Geoeye images that we all see or think of as photos on Google Earth imagery. It could have been much more fruitful if we had done that selection on site or in concert with local coffee association partners who could verify our selection of training sites.

Machine learning classifiers

The machine-learning classifiers implemented with R in this study were traditional, supervised pixel-based classifiers as opposed to object-based segmenting classifiers. The maximum likelihood segmenting classification used by TerrSet to derive training set vectors worked well for that purpose (using principal component layers as inputs), but the immature coffee class didn't survive as an identifiable pattern through that classification process. It is conceivable that a multi-resolution segmentation or pattern-recognition or image recognition-based system, such as convolutional neural networks, that can learn from chips of photos and images will be able to learn, remember, and readily identify those types of features immature

coffee features in time, just like it can recognize what is a cat, a dog, a city, or an agricultural field.

Overall there is no clear consensus on what is a good machine learning classifier for remote sensing, maybe because it's too early and there have not been quite enough adopters. On the top of the list of dominant "best classifier" characteristics, though, and perhaps mistakenly, is the theme of model accuracy. Probably every researcher involved in coffee classification that was referenced in this study arrived at different quantitative conclusions about the quality of their classifications. Each of them were working in a different region, though some were working in Central America with similar datasets to each other. At least one of these researchers had three dates of Landsat images coinciding with our own, yet the conclusions at which they arrived were dramatically different to the conclusions at which we arrived. What is noticeable is that all of them were working with different classifiers or different versions of the similar classifiers. And it's not because they're daft that the results don't yield the same answers; but because the classifiers they're using expected different inputs, input parameter ranges, or were tuned to accept a narrower selection of default parameters, or totally from another perspective that the orientation of the input data was not what the classifier was expecting. Thus, what makes a good classifier is one which can be very flexible, can work with a very large set of default values and still be able to yield good results for that wide variety of inputs (Amancio et al, 2013). Commercial remote sensing packages have not yet included many 'machine learning' classifiers in their toolsets. Most people describing these classifiers discuss them in terms of accuracy of output; our efforts in this study and previously tell us that a classifier is not necessarily best but that certain classifiers appeal to certain types of datasets and certain levels of processing of data do better for some than others. For example, principal components analyses are data dependent, however the algorithm seems to work better when data are normalized or transformed. Certain classifiers don't want to see extreme variability in their inputs (e.g. k-Nearest Neighbor seems to do be impacted by non-normalized data; and the normalization method seems to be classifier specific as well). Judging by the accuracy data presented (Table 3.5), the Maximum Likelihood Classifier (MLC) classifier had the best performance of all these efforts with Naïve_Bayes classifier coming in second position with two higher accuracies of both Sentinel 2 and Landsat 8 The worst experience (least accurate) achieved by a classifier was made by

the SVM Linear classifier, but only by 1000th of a percentage point from the Neural Network classifier. So, one thing learned it that probably a lot more classifications need to be carried out before one can say with confidence what is the best classifier of the machine learning tools in R for remotely-sensed data.

Conclusion

In this study, we analyzed relationships of texture measures with four different growth habits of coffee derived from Worldview 2, Landsat 8, Sentinel 1 & 2 images, and assessed the accuracy of classifying several land cover classes including these different coffee growth habits. We found that two pattern TerrSet texture measures, Relative Richness and Diversity (when applied to the Sentinel 1A SAR), were significant predictors of shade coffee and forest cover. GLCM measures were not found to be significant predictors of coffee growth habits. We applied machine learning classifiers (Neural network, Naïve Bayes, K-Nearest Neighbor, Random Forests and Support Vector Machines) with R to identify pasture, forest, coffee classes of shade, sun, adult, and immature and analyzed coffee map accuracy with confusion matrices. Results indicated that 80 percent accuracy is feasible with a Naïve Bayes classifier with Sentinel 2 data alone. We also found that a hybrid optical, infrared, vegetation index with effective incidence angle predictors was somewhat effective, achieving at most a 79% accuracy. No increased accuracies over previous methods were achieved. It is remains conceivable that texture can be found to be more helpful in the classification of coffee with additional work to overcome limitations.

Acknowledgements

We gratefully acknowledge the assistance from our land management partners in this research including the Tropical Agricultural Research and Higher Education Center (CATIE)’s Mesoamerican Agroenvironmental Program (MAP Norway) for their assistance in field access and data collection, Central American Commission for Environment and Development (CCAD), Trinational Commission of the Trifinio Plan (CTPT), Proyecto Trinacional Café Especial Sostenible (Protcafes), GIZ Forests and Water Program, and the US Fish and Wildlife Service.

Funding

This work was funded by NASA LCLUC Grant NNX13AC70G (K. Jones, PI) and MICITT/CONICIT in Costa Rica, and the University of Idaho Graduate Fellowship Program.

References

Amancio, D. R., Comin, C. H., Casanova, D., Travieso, G., Bruno, O. M., Rodrigues, F. A., & da Fontoura Costa, L. (2014). A Systematic Comparison of Supervised Classifiers. *PLOS ONE*, *9*(4), e94137. https://doi.org/10.1371/journal.pone.0094137

Anys, H., & He, D.-C. (1995). Evaluation of textural and multipolarization radar features for crop classification. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(5), 1170–1181. https://doi.org/10.1109/36.469481

Artiga, R. (2003). *The Case of the Trifinio Plan in the Upper Lempa: Opportunities and Challenges for the Shared Management of Central American Transnational Basins*. Unesco. Retrieved from http://webworld.unesco.org/water/wwap/pccp/cd/pdf/case_studies/the_case_of_the_trifinio_plan_in_the_upper_lempa_2.pdf

Avelino, J., Zelaya, H., Merlo, A., Pineda, A., Ordoñez, M., & Savary, S. (2006). The intensity of a coffee rust epidemic is dependent on production situations. *Ecological Modelling*, *197*(3), 431–447. https://doi.org/10.1016/j.ecolmodel.2006.03.013

Barsi, J. A., Lee, K., Kvaran, G., L. Markham, B., & A. Pedelty, J. (2014). The Spectral Response of the Landsat-8 Operational Land Imager. *Remote Sensing*, *6*, 10232–10251. https://doi.org/10.3390/rs61010232

Bivand, R. (2017). rgdal: Bindings for the "Geospatial" Data Abstraction Library (Version 1.2-15). Retrieved from https://cran.r-project.org/web/packages/rgdal/rgdal.pdf

Blackman, A., Ávalos-Sartorio, B., & Chow, J. (2012). Land cover change in agroforestry: Shade coffee in El Salvador. *Land Economics*, *88*(1), 75–101.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/10.1016/j.isprsjprs.2009.06.004

Bolanos, R. M. (2017). Producción de café es insostenible en el país. *Prensalibre*. Retrieved from http://www.prensalibre.com/economia/economia/produccion-de-cafe-es-insostenible-en-el-pais

Bolanos, S. (2007). *Using image analysis and GIS for coffee mapping*. ProQuest.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Card, D. H. (1982). Using Known Map Category Marginal Frequencies to Improve Tematic Map Accuracy. *Photogrammetric Engineering and Remote Sensing*, *48*, 431–439.

Castaneda, H. (2009). *Analysis of the spatial dynamics and drivers of forest cover change in the Lempa River Basin of El Salvador*. [Gainesville, Fla.]: University of Florida. Retrieved from http://purl.fcla.edu/fcla/etd/UFE0024235

Catalan, M. (2015). *Mapa Forestal por Tipo y Subtipo de Bosque, 2012, GUATEMALA* (Technical Report). San Salvador: INAB. Retrieved from http://www.sifgua.org.gt/Documentos/Cobertura%20Forestal/Cobertura%202012/Informe_de_Cobertura_Forestal_20_julio_15.pdf

Cathalac. (2011). *Cobertura y Uso de la Tierra de la Región del Trifinio Estudio de los anos 1986, 2001 y 2010 mediante metodos de teledeteccion*. Panama: Centro del Agua del Trópico Húmedo para América Latina y el Caribe.

CEDICAFE, & ANACAFE. (2017). Situación Nacional de la Roya del Café y Recomendaciones para su Manejo, 5.

Chen, G., & E, D. (2012). Support vector machines for cloud detection over ice-snow areas. *Geo-Spatial Information Science*. Retrieved from http://www.tandfonline.com/doi/pdf/10.1007/s11806-007-0047-7

Chuang, Y.-C. M., & Shiu, Y.-S. (2016). A Comparative Analysis of Machine Learning with WorldView-2 Pan-Sharpened Imagery for Tea Crop Mapping. *Sensors (Basel, Switzerland)*, *16*(5). https://doi.org/10.3390/s16050594

Classen, A., Peters, M. K., Ferger, S. W., Helbig-Bonitz, M., Schmack, J. M., Maassen, G., … Steffan-Dewenter, I. (2014). Complementary ecosystem services provided by pest predators and pollinators increase quantity and quality of coffee yields. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1779). https://doi.org/10.1098/rspb.2013.3148

Cordero-Sancho, S., & Sader, S. A. (2007). Spectral analysis and classification accuracy of coffee crops using Landsat and a topographic-environmental model. *International Journal of Remote Sensing*, *28*(7), 1577–1593. https://doi.org/10.1080/01431160600887680

Cruz-Bello, G. M., Eakin, H., Morales, H., & Barrera, J. F. (2011). Linking multi-temporal analysis and community consultation to evaluate the response to the impact of

Hurricane Stan in coffee areas of Chiapas, Mexico. *Natural Hazards*, *58*(1), 103–116. https://doi.org/10.1007/s11069-010-9652-0

Danilla, C. (2017). *Convolutional Neural Networks for Denoising and Classification of SAR Images*. University of Twente, Enschede, Netherlands. Retrieved from http://www.itc.nl/library/papers_2017/msc/gfm/danilla.pdf

Donald, P. F. (2004). Biodiversity Impacts of Some Agricultural Commodity Production Systems. *Conservation Biology*, *18*(1), 17–37.

Eastman, J. R. (1989). IDRISI : A geographic information system for international development. Presented at the Conference on Information Technologies for Developing Countries, California: University of Southern California.

ESA. (2010). *ESA Data Distribution Policy*. European Space Agency.

Fagan, Matthew, D., Ruth. (2015). Measurement and Monitoring of the World's Forests: A Review and Summary of Remote Sensing Technical Capability, 2009–2015. Resources for the Future. Retrieved from http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-Rpt-Measurement%20and%20Monitoring_Final.pdf

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., … others. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, *45*(2). Retrieved from http://onlinelibrary.wiley.com/doi/10.1029/2005RG000183/full

Fischer, E. F., & Victor, B. (2014). High-End Coffee and Smallholding Growers in Guatemala. *Latin American Research Review*, *49*(1), 155–177.

Geist, H. J., & Lambin, E. F. (2001). What Drives Tropical Deforestation? Retrieved June 26, 2015, from http://www.pik-potsdam.de/~luedeke/lucc4.pdf

Geist, H. J., & Lambin, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *BioScience*, *52*(2), 143.

Gomez, C., Mangeas, M., Petit, M., Corbane, C., Hamon, P., Hamon, S., … Despinoy, M. (2010). Use of high-resolution satellite imagery in an integrated model to predict the distribution of shade coffee tree hybrid zones. *Remote Sensing of Environment*, *114*(11), 2731–2744. https://doi.org/10.1016/j.rse.2010.06.007

Greenberg, R., Bichier, P., Cruz Angon, A., & Reitsma, R. (1997). BIRD POPULATIONS IN SHADE AND SUN COFFEE PLANTATIONS.pdf. *Conservation Biology*, *11*(2), 448–459.

Gross, J., & Liggs, U. (2015). Nortest: Tests for normality (Version 1.0-4) [R]. CRAN. Retrieved from https://cran.r-project.org/web/packages/nortest/nortest.pdf

Hall-Beyer, M. (2017). GLCM Texture: A Tutorial v. 3.0 March 2017. Retrieved from http://prism.ucalgary.ca//handle/1880/51900

Haralick, R., Dinstein, I., & Shanmugam, K. (1973). Textural Features for Image Classification. *IEEE Transactions On Systems, Man and Cybernetics*, *SMC-3*(6), 610–621.

Hecht, S. B., & Saatchi, S. S. (2007). Globalization and forest resurgence: Changes in forest cover in El Salvador. *Bioscience*, *57*(8), 663–672.

Herold, N. D., Haack, B. N., & Solomon, E. (2004). An evaluation of radar texture for land use/cover extraction in varied landscapes. *International Journal of Applied Earth Observation and Geoinformation*, *5*, 113–128. https://doi.org/10.1016/j.jag.2004.01.005

Holben, B. N., & Justice, C. O. (1980). The Topographic Effect on Spectral Response from Nadir-Pointing Sensors. *Photogrammetric Engineering & Remote Sensing*, *46*(9), 1191–1200.

Jezeer, R. E. (n.d.). Shaded Coffee and Cocoa – Double Dividend for Biodiversity and Small-scale Farmers. Retrieved October 8, 2017, from https://www.researchgate.net/publication/316827107_Shaded_Coffee_and_Cocoa_-_Double_Dividend_for_Biodiversity_and_Small-scale_Farmers

Jha, S., Bacon, C. M., Philpott, S., Mendez, E., Laderach, P., & Rice, R. (2014). Shade Coffee: Update on a Disappearing Refuge for Biodiversity. *Bioscience*, *64*(5), 416–28.

Jimenez, A. J. (2014). *Final Report of ES Forest Map 2011, July 2014*. San Salvador.

Kailath, T. (1967). The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transactions on Communication Technology*, *15*(1), 52–60. https://doi.org/10.1109/TCOM.1967.1089532

Karakizi, C., Oikonomou, M., & Karantzalos, K. (2016). Vineyard Detection and Vine Variety Discrimination from Very High Resolution Satellite Data. *Remote Sensing*, *8*(3), 235. https://doi.org/10.3390/rs8030235

Kleinschroth, F., & Healey, J. R. (2017). Impacts of logging roads on tropical forests. *Biotropica*, *49*(5), 620–635. https://doi.org/10.1111/btp.12462

Kuhn, M. (2014). *Applied Predictive Modeling in R*. Groton, CT. Retrieved from http://static.squarespace.com/static/51156277e4b0b8b2ffe11c00/t/53ad86e5e4b0b52e4e71cfab/1403881189332/Applied_Predictive_Modeling_in_R.pdf

Langford, M., & Bell, W. (1997). Land cover mapping in a tropical hillsides environment: a case study in the Cauca region of Colombia. *International Journal of Remote Sensing*, *18*(6), 1289–1306.

Lelong, C., Alexandre, C., & Dupuy, S. (2014). Discrimination of tropical agroforestry systems in very high resolution satellite imagery using object-based hierarchical classification: A case-study in Cameroon. *South-Eastern European Journal of Earth Observation and Geomatics*, *3*(2S), 255–258.

Lelong, C. C. D., & Thong-Chane, A. (2003). Application of textural analysis on very high resolution panchromatic images to map coffee orchards in Uganda. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)* (Vol. 2, pp. 1007–1009 vol.2). https://doi.org/10.1109/IGARSS.2003.1293994

Leutner, B., & Horning, N. (n.d.). RStoolbox: Tools for Remote Sensing Data Analysis (Version R Package version 0.1.9). Retrieved from https://cran.r-project.org/web/packages/RStoolbox/index.html

Li, F., Jupp, D. L. B., Reddy, S., Lymburner, L., Mueller, N., Tan, P., & Islam, A. (2010). An Evaluation of the Use of Atmospheric and BRDF Correction to Standardize Landsat Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(3), 257–270. https://doi.org/10.1109/JSTARS.2010.2042281

Lu, D., Li, G., Moran, E., Dutra, L., & Batistella, M. (2014). The roles of textural images in improving land-cover classification in the Brazilian Amazon. *International Journal of Remote Sensing*, *35*(24), 8188–8207. https://doi.org/10.1080/01431161.2014.980920

MARN. (April 20016). MARN lanza el Plan Nacional de Restauración y Reforestación | MARN | Ministerio de Medio Ambiente y Recursos Naturales [Government]. Retrieved July 11, 2016, from http://www.marn.gob.sv/marn-lanza-el-plan-nacional-de-restauracion-y-reforestacion/

Martignoni, M. (2011). Land use and cover classification using airborne MASTER and spaceborne GeoEye-1 sensors: Focus on coffee-banana agroforestry systems near Turrialba. *Costa Rica*. Retrieved from http://agroforestbanana.org/files/documentos/M_Martignoni_thesis_final.pdf

Martínez-Verduzco, G. C., Galeana-Pizaña, J. M., & Cruz-Bello, G. M. (2012). Coupling Community Mapping and supervised classification to discriminate Shade coffee from Natural vegetation. *Applied Geography, 34*, 1–9. https://doi.org/10.1016/j.apgeog.2011.10.001

Menendez Martinez, A. F. (2015). *Sustainable Coffee Production in the Trifinio Region, Guatemala, Honduras, El Salvador* (Annual). Esquipulas: ICP. Retrieved from http://www.coffee-partners.org/files/uploads/documents/Downloads_Trifinio/1503%20ICP%20Trifinio%20-%20Annual%20Report.pdf

Monmonier, M. S. (n.d.). Measures of Pattern Complexity for Choropleth Maps. *The American Cartographer*, *1*(2), 159–169.

Montenegro, R., & Atwood, D. (2010). *Identification of Shade Coffee Using Optical / SAR Data Fusion*. Poster presented at the 2010 AGU Meeting of the Americas, Foz do Iguacu, Brazil. Retrieved from http://www.agu.org/meetings/ja10/pdf/AGU_JA10_Program.pdf

Mukashema, A., Veldkamp, A., & Vrieling, A. (2014). Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network. *International Journal of Applied Earth Observation and Geoinformation*, *33*, 331–340. https://doi.org/10.1016/j.jag.2014.05.005

Muñoz Brenes, C. L. (2017, April 24). *Crossing the Line on Governance: Evaluating the Impact of National and Transboundary Protected Areas on Land Cover Outcomes in Central America*. University of Idaho, Moscow.

Nagendra, H., Southworth, J., & Tucker, C. (2003). Accessibility as a determinant of landscape transformation in western Honduras: linking pattern and process. *Landscape Ecology*, *18*(2), 141–158.

Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, *29*(1), 51–59. https://doi.org/10.1016/0031-3203(95)00067-4

Ortega-Huerta, M. A., Komar, O., Price, K. P., & Ventura, H. J. (2012). Mapping coffee plantations with Landsat imagery: an example from El Salvador. *International Journal of Remote Sensing*, *33*(1), 220–242. https://doi.org/10.1080/01431161.2011.591442

Ozdarici Ok, A., & Akyurek, Z. (2012). A segment-based approach to classify agricultural lands by using multi-temporal optical and microwave data. *International Journal of Remote Sensing*, *33*(22), 7184–7204. https://doi.org/10.1080/01431161.2012.700423

Ozdemir, I., & Karnieli, A. (2011). Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation*, *13*(5), 701–710. https://doi.org/10.1016/j.jag.2011.05.006

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., & Thirion, B. (2011). Scikit-learn: Machine Learning in Python, *12*, 6.

Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing*, *50*(1), 144–154. https://doi.org/10.1080/22797254.2017.1299557

Rahman, M. M. (2010). Mapping tropical forest and deforestation using synthetic aperture radar (SAR) images. *Applied Geomatics*, *2*, 113–121.

REDDCCADGIZ. (2011). *Sistematización de Experiencias Regionales, Análisis de la Situación Actual y Propuesta para la Operativización de un Sistema de Monitoreo de Bosques Multinivel para los Países de Centroamérica y República Dominicana* (p. 86). SanSalvador. Retrieved from http://www.reddccadgiz.org/documentos/doc_1148286616.pdf

Rice, R. (2010). The ecological benefits of shade-grown coffee: the case for going bird friendly. *Smithsonian Migratory Bird Center, Washington*. Retrieved from http://nationalzoo.si.edu/scbi/migratorybirds/coffee/bird_friendly/Eco-Report.pdf

Richards, M. B., & Méndez, V. E. (2014). Interactions between Carbon Sequestration and Shade Tree Diversity in a Smallholder Coffee Cooperative in El Salvador. *Conservation Biology*, *28*(2), 489–497. https://doi.org/10.1111/cobi.12181

Rouse, J., Haas, R., Schell, J., & Deering, D. (1973). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation* (Progress Report) (p. 112). Greenbelt, MD: NASA GSFC.

Rueda, X., Thomas, N. E., & Lambin, E. F. (2015). Eco-certification and coffee cultivation enhance tree cover and forest connectivity in the Colombian coffee landscapes. *Regional Environmental Change*, *15*(1), 25–33. https://doi.org/10.1007/s10113-014-0607-y

Schlesinger, P., Muñoz Brenes, C. L., Jones, K. W., & Vierling, L. A. (2017). The Trifinio Region: a case study of transboundary forest change in Central America. *Journal of Land Use Science*, *12*(1), 36–54. https://doi.org/10.1080/1747423X.2016.1261948

Schmitt-Harsh, M. (2013). Landscape change in Guatemala: Driving forces of forest and coffee agroforest expansion and contraction from 1990 to 2010. *Applied Geography*, *40*, 40–50. https://doi.org/10.1016/j.apgeog.2013.01.007

See, L., Laso Bayas, J. C., Schepaschenko, D., Perger, C., Dresel, C., Maus, V., … Fritz, S. (2017). LACO-Wiki: A New Online Land Cover Validation Tool Demonstrated Using GlobeLand30 for Kenya. *Remote Sensing*, *9*(7), 754. https://doi.org/10.3390/rs9070754

Sesnie, S. E., Finegan, B., Gessler, P. E., Thessler, S., Ramos Bendana, Z., & Smith, A. M. S. (2010). The multispectral separability of Costa Rican rainforest types with support vector machines and Random Forest decision trees. *International Journal of Remote Sensing*, *31*(11), 2885–2909. https://doi.org/10.1080/01431160903140803

Sesnie, S. E., Hagell, S. E., Otterstrom, S. M., Chambers, C. L., & Dickson, B. G. (2008). SRTM-DEM and landsat ETM+ data for mapping tropical dry forest cover and biodiversity assessment in Nicaragua. *Rev. Geogr. Acad*, *2*, 53–65.

Smith, E. S. (2010). The Evolution of Coffee Markets for Sustainable Development: A Honduran Cooperative's Experience With Fair Trade. Retrieved from https://www.csuchico.edu/anth/pdf/Erin%20Smith%20thesis.pdf

Soh, L.-K., & Tsatsoulis, C. (1999). Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices, *47*, 17.

Stiftung, H. R. N. (n.d.). Trifinio Regional Program. Retrieved from https://www.hrnstiftung.org/projekte/trifinio-regional-program

Tang, Z., Qi, F., Zhou, Y., Pan, F., & Zhou, J. (2015). Tea Leaves Classification Based on Texture Analysis. In *Proceedings of the 2015 Chinese Intelligent Automation Conference* (pp. 353–360). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-46469-4_37

Tonye, E. (2002). Evaluation of speckle filtering and texture analysis methods for land cover classification from SAR images. *International Journal of Remote Sensing*, *23*(9), 1895–1925.

Turner, M. G. (1989). Landscape Ecology: The Effect of Pattern on Process. *Annu. Rev. Ecol. Syst*, *20*, 171–197.

USGS. (2014). *Landsat Surface Reflectance Climate Data Records*. Retrieved from https://pubs.usgs.gov/fs/2013/3117/pdf/fs2013-3117.pdf

USGS. (2017, October). Landsat 4-7 Surface Reflectance (LEDAPS) Product Guide. USGS. Retrieved from https://landsat.usgs.gov/sites/default/files/documents/ledaps_product_guide.pdf

USGS. (n.d.). Solar Illumination and Sensor Viewing Angle Coefficient Files. Retrieved November 15, 2017, from https://landsat.usgs.gov/solar-illumination-and-sensor-viewing-angle-coefficient-file

van Rossum, G. (1995). *Python Tutorial* (Technical report CS-R9526). Amsterdam: Centrum voor Wikunde en Informatica (CWI). Retrieved from http://www.oalib.com/references/12583079

Whiteside, T., & Bartolo, R. (2014). Vegetation map for Magela Creek floodplain using WorldView-2 multispectral image data. Retrieved from http://www.environment.gov.au/system/files/resources/448c1688-77e6-4cc2-a5b6-af1280d75048/files/ir628_3.pdf

Zhang, X., Cui, J., Wang, W., & Lin, C. (2017). A Study for Texture Feature Extraction of High-Resolution Satellite Images Based on a Direction Measure and Gray Level Co-Occurrence Matrix Fusion Algorithm. *Sensors*, *17*(7). https://doi.org/10.3390/s17071474

Zhou, T., Pan, J., Han, T., & Wei, S. (2017). Planting area extraction of winter wheat based on multi-temporal SAR data and optical imagery. *Transactions of the Chinese Society of Agricultural Engineering*, *33*(10), 215–221.

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

CHAPTER 4: CONCLUSION

In this dissertation, I investigated changes in forested land cover and land use across the Trifinio Region over a 30-year period, and studied how these changes may have differed from elsewhere in Central America. I tested whether regional forest transitions and forest resurgence were related to proximity to PAs, agricultural fields, pastures, or shaded coffee production. I specifically worked with Landsat 8, Sentinel 1, Sentinel 2, and Worldview 2 images to test whether classifications derived from these new sensor data would show different accuracies than already reported use with machine learning classifiers: Neural network, Naïve Bayes, K-Nearest Neighbor, Random Forests and Support Vector Machines, and Maximum Likelihood.

Land cover and land use changes initially measured using secondary data, and subsequently databases of percent greenness, determined that the types and drivers of transitions in the Trifinio Region over the last 30 years were no different than those affecting other tropical areas. We modeled deforestation with expected drivers and found relationships between population density and extreme deforestation. We also found correlations between road density and deforestation. While transitions of greenness pointed to temporal increases over the 30-year period, we did not find a signal of forest resurgence corresponding with the reports of earlier investigators. It is conceivable that the percent greenness transitions methods we used could have masked the changes we wanted to find; for example, while broadleaved forest has higher greenness when growing rapidly, coniferous forest has lower values, and we were not able to separate these data into forest types like this. Forest transitions in the study period did not correlate in any large way with abandoned agricultural fields and pastures, nor with proximity to areas of shaded coffee production. We did find considerably more samples of increasing slopes in and near PAs within the Trifinio Region relative to areas more distant from protected areas, and this is a finding that deserves future investigation. It is conceivable that the dry season timing of the time series of percent greenness data impacted the outcome of this study.

We found that two texture measures (relative richness and diversity) are responsive to patterns of forest cover and coffee production land uses. Landsat 8 images offered the best data for mapping these neotropical forests and their coffee land uses. The new Sentinel 1

radar sensor provide data regardless of cloud cover, but did not provide any additional power to detect forest or coffee relative to optical sensors. In the future, it would be recommended to re-explore this question with L-band systems that penetrate deeper into the vegetative canopy than the C-band of Sentinel 1. We found little evidence that the Sentinel 1 active sensor images were suitable to identify forest or coffee. The influence of classifying land cover using radar was not affected by our choice of classifier. The Maximum Likelihood classifier offered the best accuracy among tested classifiers. We achieved medium accuracy using a hybrid approach that had been identified a decade ago using older Landsat images and effective incidence angles to offset the impact of the angle of solar reflectance. Future activities should seek to conduct similar assessments with multi-polarized radar data, as these have reportedly offered increased abilities to identify surfaces. Better onsite connections, working directly with local farmers and coffee cooperatives, and local researchers could improve the quality of training site materials. This work did blaze a trail for those who will come later, as few have yet conducted such work in Central America, let alone the Trifinio Region. The need for further education and connection of land use science with patterns of land cover changes remains.

Uncertainties and Caveats

Potential Uncertainties of the Overall Work

There are a number of relevant uncertainties that could be mentioned (organized by relevance to Chapter), including:

Chapter 1: 1, 3, 4, 6

Chapter 2: 1, 2, 3, 4, 5, 6, 7, 11, 12, 13, 14

Chapter 3: 2, 3, 4, 6, 7, 8, 9, 10, 11

1. Landsat 5 viewing geometry changes.

This study tried to use Landsat 5 as much as possible, because it has been the most consistent satellite sensor over time. However, a recent review of the 27-year archive of this satellite showed there have been small but noticeable inconsistencies in the viewing geometry

due to a combination of orbital changes and reflectance anisotropy. These impacted the sensor's NDVI values over time, with the overpass time changing about an hour, and solar zenith by more than 10 degrees (Zhang and Roy, 2016). While this certainly would have impacted the current study, it offers another reason for the use of quintiles as a conservative method to quantify change.

2. Using NDVI as a proxy for landcover or forest cover change.

As my initial efforts to map forest cover in the Trifinio Region were spectrally confused by shade coffee production, as noted in Chapter 2, I decided to use NDVI as a proxy for forest cover. However, NDVI as a measure of the relative degree of greenness, does not offer a result the same as classical remotely-sensed maps (with supervised classifications and hard boundaries). Land cover (e.g. broadleaf, conifer, and mixed forest) and land use (coffee, agriculture, pasture, urban, and others) change for each of the seven temporal epochs were desired for comparison over time. My conclusion for the initial mapping failed because it not did not separate the predicted forest cover layer into forest and coffee classes (e.g. a sun coffee, a shade coffee, an immature coffee, a mature coffee) because of the spectral confusion between shade coffee and forest cover. A comparison of the predicted forest layer with very high-resolution data showed that the forest cover predicted in many locations was in fact natural vegetation-based coffee shading components, as well as coffee itself.

3. Lack of illumination correction.

None of the satellite data used in Chapter 2 and Chapter 3 were corrected for illumination (impacts of topographic slope and aspect) (with exception of a single test in Chapter 3 to include illumination correction created with different parts of the TIRS spectrum (via bands 10 and 11).

4. Contentious issue of signal saturation (for tropical evergreen of moderate to high biomass) (Yengoh et al., 2016).

Some scientists believe that NDVI saturation is a potential uncertainty with tropical evergreen areas of moderate to high biomass. If this issue is true, then this was a potential issue for study in several of the protected areas, especially in the Celaque National Park in Honduras, where no significant regrowth signal was found that had been reported by other

researchers 20 years earlier. However, Yengoh et al., (2016) argue that in fact saturation does not occur, and that instead photosynthesis is at a maximum because no more light can be absorbed.

5. Temporal transition class assignments.

Conservative class labels (20% categories - called quintiles) were assigned purposefully to the land cover transitions to avoid any chance of creating change just because a small threshold had been surpassed. Three types of changes were permitted (increasing, flat, and decreasing), with 15 and 30-year periods containing the entire study period and the most recent 15-year period. Different results may have occurred if we had considered the period of 1986-2001, as multiple directions could have been encountered.

6. Dry season timing of our study.

The dry season is a difficult time to see forest cover (as most broadleaf trees are in leaf-off). There was no other option as the cloud cover was extensive in all other images of the region.

7. Single polarization of RADAR.

Single polarized RADAR (was the only type available for during the study period. Dual polarized data are available now, and these data together been proven to achieve much better success at recognizing ground cover.

8. Lack of onsite components (no onsite field work).

As a result of security issues, we did not go to the field in the study site. However, it would have been more fruitful to have had a better feeling for the vegetation types and to have gained coordinate data via GPS for shade and sun coffee training sites.

9. Use of Google Earth for validation of validation maps.

Google Earth is not a completely reliable source for mapping. Google Earth mapped data was used as a method of last resort for verifying the national forest maps used as validation maps for accuracy assessment.

10. Mismatch between dates of validation maps.

There was a mismatch of years (as many as 5 years) among the validation map sources that were joined for the study subset to measure accuracy in Chapter 3.

11. Mismatch between satellite image dates.

There was an unavoidable mismatch between one satellite image acquisition date that likely was impacted by a small green-up due to rainfall in the study region, but there was no alternate image to use because of cloud cover.

12. Large sample size affected p-values.

Texture predictors underwent transformations of square root, logarithm, and cube root and subjected to Logistic Regressions to seek significant correlates in R. The sample size was 72,000 (10% of the 721,000 stratified random point samples within Trifinio Region. It is quite likely that this impacted our significant finding of texture measures being statistically significant predictors of shade coffee.

13. Very small, but statistically-significant p-values.

Very small, but statistically-significant p-values almost certainly impacted our outcome, probably due to the very large sample size. For example, shade coffee and diversity have a lower Confidence Interval bound of 0.0024, therefore the increase in likelihood from a one-unit increase in diversity would have only impacted our observation likelihood by 1.0002430 times.

14. Flat slope transitions has several positions higher and lower.

As flat slope transition has several potential positions (e.g. 5555555, 4444444, 3333333, and 2222222), as well as the type that oscillates up and down by a single quintile over the seven epochs. It would be important to ascertain which type of class is apparent for any one region, and additionally to determine what happens to the flat slope at other parts of the year, by comparing its location spatially and spectrally with another method of monitoring NDVI (e.g. MODIS or GIMMS).

What does this all mean? How does it relate to LCLUC drivers?

My dissertation contributes to the science of LCLUC in the development of knowledge about changes in land cover and in response to drivers of land use transitions in a transboundary region of Central America. The NDVI trajectories (in the form of temporal transition classes) function as indicators for land cover degradation and vegetative recovery (and soil color changes) (Tovar, 2012). Accommodating some of the simple uncertainties (illumination correction, improving onsite knowledge, adding dual polarization to the RADAR band) should improve our understanding of the loses and gains. Landsat NDVI is already a useful tool to understand local processes and alert management to high level changes, but though often inconsistent (Cui et al, 2013; Zhang & Roy, 2015). Reducing its uncertainties will facilitate better understanding of forest loss drivers.

Alternative data and measures

Wilson and Sader (2002) reviewed and compared other methods in a temperate forest exercise (at a hardwood site, leaf-on, and during the summer in the State of Maine) and included the Normalized Difference Moisture Index (NDMI) which uses the mid-infrared (MIR) band instead of the near-infrared, and found it superior to NDVI in a series of classifications of multi-layer change maps, investigating interval lengths between changes. The MIR band, an absorber of water, had reduced reflectance. The number of bands used and the length of the interval in the change detection, were important, with the highest six interval change created with seven dates of images 2-3 years apart being the most accurate. This compares well with my effort, except that the length of time between the limits of the intervals are longer in the Trifinio work, mostly 5, but as much as 8 years. Alternative measures include tasseled cap transformations (TCT) that can convert Landsat image layers into a set of wetness, greenness, and brightness layers that can be used to identify vegetation state, moisture, and degree of soil background & senescence. Two alternative data sets, recommended by literature are considered the "best" data for long term analyses of NDVI (MODIS & GIMMS3g) (Yengoh et al, 2014; Mbow et al, 2015), however neither of these are likely very useful for sensing smallholder Central American agricultural assessment due to pixel resolution (250 meters & 1 kilometer, respectively), but for wide-area assessment time series work, these could be applied and compared with the Landsat short time series result.

Land cover restoration and land use management efforts have been carried out elsewhere using NDVI time-series analyses. A comprehensive publication on the global experience with NDVI with useful examples of sub-global land use management via remote sensing can be found in Appendix A of Yengoh et al. (2015) "Inventory of Some Global and Sub-global Remote Sensing-Based Land Degradation Assessments". Topics of relevance to the Trifinio Region include drought, vegetation dynamics, ecosystem resilience, land use and land cover change, and drivers of degradation. Satellite image products used included Advanced Very High-Resolution Radiometer (AVHRR), Global Inventory for Mapping and Modeling Studies (GIMMS), Moderate-Resolution Imaging Spectroradiometer (MODIS), Landsat, and Pathfinder AVHRR Land (PAL). Three of these efforts used Landsat data, and included topics of LCLUC, Soil Organic Carbon and Salinization, Desertification, and Vegetation Burning and Recovery. LCLUC was assessed with two images of southern Italy and the NDVI differencing approach (Mancini et al, 2014). In a study on desertification, Xu et al (2009) examined the Ordos region of Chinese Inner Mongolia using 21 Landsat MSS and TM images from 1980-2000 and NDVI and the Moving Standard Deviation Index (MSDI) to characterize desertification in 5 grades. The project developed rule-sets (thresholds to assign differing levels of NDVI, MSDI, and Albedo) to differing desertification stages (non, low, medium, high, and severe) for use with a decision tree. In a study on fire severity, Diaz-Delgado et al (2003) used summed Landsat TM NDVI and fire severity classes to monitor regeneration after fire; they built a new variable of avg NDVI (in burned parcel by severity) divided by avg NDVI of a control location, to better understand the impact of fire severity on regeneration.

Potential next steps

Chapter 1

A new 2016 map for Trifinio needs to be created to compare land use change between CATHALAC and current day. These data must be corrected for illumination and can overcome the coffee spectral confusion issue using brute force cluster-busting) and validation data collected by an onsite field component, perhaps with the assistance of PROTCAFE or CUNORI.

Chapter 2

There is a need to follow up on areas with questions: including where and why are there thousands of increasing trend slopes in protected areas inside Trifinio, and why were there almost none outside of Trifinio; and what do the flat trends mean? Are they areas of degraded grasslands, irrigated agriculture, or perennial cropping? As well, a classical remote sensing approach should be sought with the Celaque National Park area for the years of the Nagendra et al (2003) study to replicate its finding and determine if these trees remained growing. It would useful to compare the slope transition classes with the tasseled cap wetness (Lea et al, 2003) procedure to verify to clarify our understanding of the transitions.

Chapter 3

The shade coffee-texture work should be replicated in Turrialba, Costa Rica, using a variety of large texture windows identified as useful by other researchers, and improved with an onsite component perhaps connected with Gamma, CIRAD, and Masters students to create better training and validation data for a coffee identification activity? convolutional neural network (CNN)-based model and classification should be constructed and tested to classify immature, adult, and shade coffees in Turrialba.

A shade coffee and immature coffee database needs to be created as does an image library for a convolutional neural network test, though it is conceivable that this can be done with Google Earth images, and a regional GIS-based coffee inventory.

A study of how object-based classifiers compare with CNN-based classifiers could be very productive.

Lastly increasing technical capacity building in remote sensing and geographic processes can help to overcome barriers in the Trifinio Region. Still additional barriers include the identification of a regional team, lab space, and support to carry on key GIS and remote sensing work and in service to the Trifinio Region. Development of a financial proposal to improve technical capacity to the Region to carry out citizen science, and to improve knowledge management (e.g. information sharing) could likely be put together with the team at CUNORI and be supported by Trifinio's SINTET information management system.

References

Cui, X.; Gibbes, C.; Southworth, J.; Waylen, P. Using Remote Sensing to Quantify Vegetation Change and Ecological Resilience in a Semi-Arid System. *Land*, 2013, 2, 108-130.

Johnson, D. H. (1999). The Insignificance of Statistical Significance Testing. *The Journal of Wildlife Management*, 63(3), 763–772. https://doi.org/10.2307/3802789

Lea, R.; Blodgett, C.; Diamond, D and Schanta, M. 2003. Using the tasseled cap transformation to identify change in the Missouri Ozark forests, MoRAP Project: Forest Change Detection.

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, 105(2), 142–154. https://doi.org/10.1016/j.rse.2006.06.018

Mancino G, Nolè A, Ripullone F, Ferrara A (2014) Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy. *iForest-Biogeosci Forestry* 7(2):76.

Mbow, C., Brandt, M., Ouedraogo, I., de Leeuw, J., & Marshall, M. (2015). What Four Decades of Earth Observation Tell Us about Land Degradation in the Sahel? *Remote Sensing*, 7(4). https://doi.org/10.3390/rs70404048

Tovar, C. L. Meneses., "NDVI as indicator of Degradation," *Unasylva*, 62. 238. 2012. Available at: www.fao.org/docrep/o 15/i2560e/i2560e07.pdf; accessed on 14-03-2014.

Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, 80(3), 385–396. https://doi.org/10.1016/S0034-4257(01)00318-2

Xu, D., Kang, X., Qiu, D., Zhuang, D., & Pan, J. (2009). Quantitative Assessment of Desertification Using Landsat Data on a Regional Scale – A Case Study in the Ordos Plateau, China. *Sensors*, 9(3). https://doi.org/10.3390/s90301738

Yengoh, G. T., Dent, D., Olsson, L., Tengberg, A. E., & Tucker, C. J. (2014). *The use of the Normalized Difference Vegetation Index (NDVI) to assess land degradation at multiple scales: a review of the current status, future trends, and practical considerations* (p. 80). Lund, Sweden: Lund University Centre for Sustainability Studies - LUCSUS. Retrieved from http://www.stapgef.org/stap/wp-

content/uploads/2015/05/Final-report-The-use-of-NDVI-to-assess-land-degradation-G.-Yengoh-et-al.pdf

Yengoh G.T., Dent D., Olsson L., Tengberg A.E., Tucker C.J. (2015) Limits to the Use of NDVI in Land Degradation Assessment. In: Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales. SpringerBriefs in Environmental Science. Springer, Cham.

Zhang, H. K., & Roy, D. P. (2016). Landsat 5 Thematic Mapper reflectance and NDVI 27-year time series inconsistencies due to satellite orbit change. *Remote Sensing of Environment*, 186(Supplement C), 217–233. https://doi.org/10.1016/j.rse.2016.08.022

Appendix A

Intermediate Results

Table A.1. Classes Mapped by Worldview 2 Segmentation and Classification

1	Pasture
2	Shade Coffee
3	Sun Coffee
4	Cloud1 bright white tops
5	Cloud2 lower, hazy, mixed shadow)
6	Forest
7	Mature Coffee
8	Immature Coffee
9	Shadow
10	Rural (housing/roofs)
11	Urban (housing/streets, roofs)
12	Shrubs
13	Burned-pasture-woodland
14	Field Water
15	River / Lake Water
16	White Stuff (bare soil or sand)
17	Tree Crown
18	Asphalt Road
19	Dirt Road
20	Wet Pasture

Table A.2. Filter Type and Size Evaluation

Speckle filter type and size was determined by comparing pre-filtering and post filtering scores of the Mean, Sigma, and Equivalent Number of Looks (ENL) values implementing the rule that the chosen filter should reduce the speckle but produce post-filtering values that don't overly compromise the dataset. In this case, we chose to use a filter that Despeckles (D), didn't Reduce (R), because reducing adaptive filters assumed there is a need to preserve edges.

MEAN	Reduce/	Kernel				MEAN DIFFERENCE (POST-FILTERING)				BEST SCORE
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.1244	0.1244	0.1244	0.1244	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
Median	R	0.1171	0.1103	0.1061	0.1033	-0.0073	-0.0141	-0.0183	-0.0211	**(larger value**
Frost	D	0.1236	0.1228	0.1229	0.1232	-0.0008	-0.0016	-0.0015	-0.0012	**= higher**
Gamma Map	R	0.1243	0.1242	0.124	0.124	**-0.0001**	**-0.0002**	-0.0004	-0.0004	**quality)**
Lee	D	0.1243	0.1242	0.124	0.124	**-0.0001**	**-0.0002**	-0.0004	-0.0004	**Lee**
Lee Sigma	D	XXXX	0.1258	0.1245	0.1234	XXXX	**0.0014**	**0.0001**	**-0.0010**	
IDAN50	R	0.1087				-0.0157				
IDAN75	R	0.1067				-0.0177				
IDAN100	R	0.1055				-0.0189				
RefinedLee	R	0.1176				-0.0068				
SIGMA										
	Reduce/	Kernel				SIGMA DIFFERENCE (POST-FILTERING)				
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.144	0.116	0.0999	0.893	-0.0421	-0.0701	-0.0862	0.7069	
Median	R	0.1227	0.0866	0.0712	0.0622	-0.0634	**-0.0995**	**-0.1149**	-0.1239	**(smaller value**
Frost	D	0.1449	0.1573	0.1629	0.164	-0.0412	-0.0288	-0.0232	-0.0221	**= higher**
Gamma Map	R	0.1435	0.1141	0.1055	0.0943	-0.0426	-0.0720	-0.0806	**-0.0918**	**quality)**
Lee	D	0.1437	0.1158	0.1015	0.1038	-0.0424	-0.0703	-0.0846	-0.0823	**Lee**
Lee Sigma	D	XXXX	0.1769	0.5112	0.172	XXXX	-0.0092	**0.3251**	-0.0141	
IDAN50	R	0.089				**-0.0971**				
IDAN75	R	0.0852				**-0.1009**				
IDAN100	R	0.083				**-0.1031**				
RefinedLee	R	0.1168				-0.0693				
ENL										
	Reduce/	Kernel				ENL DIFFERENCE (POST-FILTERING)				
Type	Despeckle	3x3	5x5	7x7	9x9	3x3	5x5	7x7	9x9	3x3
Boxcar	R	0.7459	1.1492	1.549	1.9375	0.2995	0.7028	**1.1026**	**1.4911**	
Median	R	0.9115	1.6215	2.2172	2.7564	0.4651	**1.1751**	**1.7708**	**2.31**	**(larger value**
Frost	D	0.7272	0.6095	0.5692	0.5642	0.2808	0.1631	0.1228	0.1178	**= higher**
Gamma Map	R	0.75	1.1819	1.3772	1.7128	0.3036	0.7355	0.9308	**1.2664**	**quality)**
Lee	D	0.7483	1.1502	1.4934	1.4273	0.3019	0.7038	**1.047**	**0.9809**	**Lee**
Lee Sigma	D	XXXX	0.5054	0.5112	0.5147	XXXX	0.059	0.0648	0.0683	
IDAN50	R	1.4922				**1.0458**				
IDAN75	R	1.568				**1.1216**				
IDAN100	R	1.6165				**1.1701**				
RefinedLee	R	1.0147				0.5683				

Table A.3. Landsat 8 Correlation Matrix

Optical Band Correlation/Correlation Matrices per image type. Landsat 8 bands 1-4, 6 & 7 show significant correlation, while bands 5, 9, 10, and 11 are not (Table A.3).

	B1	B2	B3	B4	B5	B6	B7	B9
B1	1	**0.99**	**0.96**	**0.95**	0.12	**0.82**	**0.87**	-0.2
B2	0.99	1	**0.97**	**0.97**	0.13	**0.84**	**0.9**	-0.17
B3	0.96	0.97	1	**0.97**	0.27	**0.85**	**0.88**	-0.12
B4	0.95	0.97	0.97	1	0.13	**0.9**	**0.93**	-0.11
B5	0.12	0.13	0.27	0.13	1	0.24	0.1	0.04
B6	0.82	0.84	0.85	0.9	0.24	1	**0.96**	-0.11
B7	0.87	0.9	0.88	0.93	0.1	0.96	1	-0.09
B9	-0.2	-0.17	-0.12	-0.11	0.04	-0.11	-0.09	1

Table A.4. Sentinel 2a Band Correlation Matrix

Sentinel 2A bands 2,3, and 4 are highly correlated while band 8 is not.

	B2	B2	B4	B8
B2	1	**0.97**	**0.96**	0.1
B3	**0.97**	1	**0.95**	0.24
B4	**0.96**	**0.95**	1	0.07
B8	0.1	0.24	0.07	1

Table A.5. Worldview 2 Correlation Matrix

Worldview 2 Correlation Frequency Matrix (by band 1-8).

	B1	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**	**B8**
B1		8	8	8	2	0	0	0
B2			8	8	5	0	0	0
B3				8	8	2	0	0
B4					8	0	0	0
B5						0	0	0
B6							8	8
B7								8

The table values indicate the frequency in number of images (of the eight adjacent images research tiles) where there was a Pearson's correlation value > 0.9.

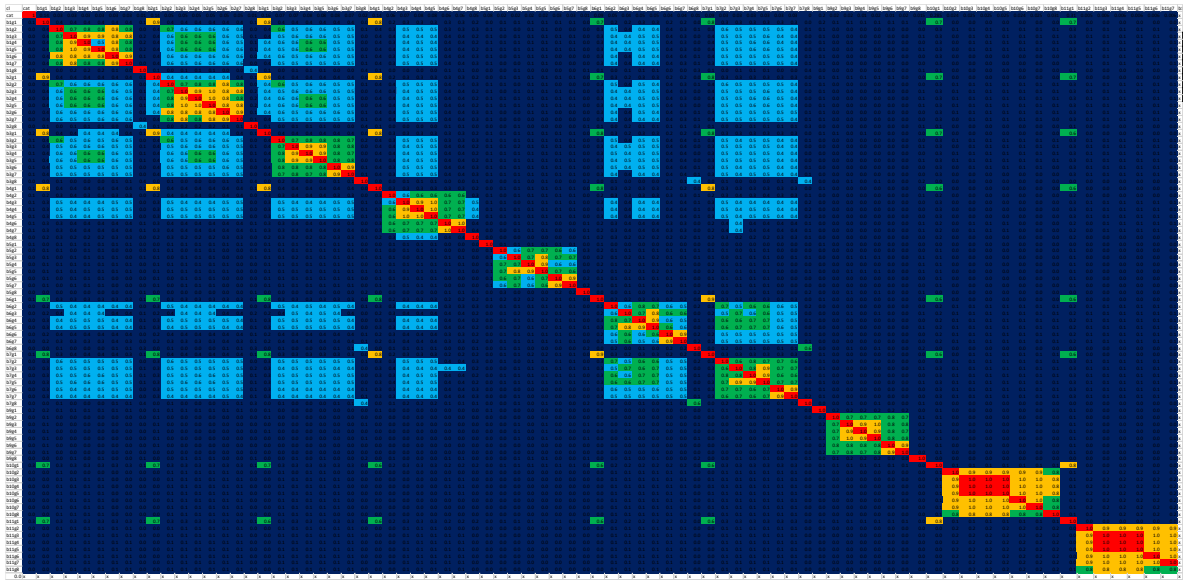
Table A.6. Principal Components Analysis Results (in percent variance)

This table shows that use of the first two components would be sufficient to capture 98-99% of the image variance. In this case, components C1 and C2 were used in segmentation of the Worldview 2 tiles to derive vector training sites.

Worldview 2	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8
4923	82.397869	16.244251	0.695184	0.400008	0.188016	0.035136	0.027354	0.012182
4924	82.550595	16.18719	0.57614	0.385112	0.242325	0.02601	0.02212	0.010508
4926	85.451593	13.350178	0.771027	0.203482	0.158698	0.029699	0.02258	0.012744
4827	85.451593	13.350178	0.771027	0.203482	0.158698	0.029699	0.02258	0.012744
4942	84.610842	14.114551	0.675077	0.325967	0.206324	0.031907	0.025721	0.00961
4943	84.720749	13.974285	0.650967	0.369879	0.19003	0.043508	0.028053	0.022531
4944	84.38651	14.220015	0.742006	0.39564	0.189132	0.030961	0.024262	0.011471
4945	84.389748	14.136968	0.834962	0.365119	0.206935	0.029234	0.025257	0.011778
LANDSAT 8	99.999871	0.000099	0.00002	0.000012	4.93E-07	4.50E-07	1.66E-07	4.75E-08
SENTINEL 2	67.516058	30.519634	1.838815	0.125487				

Table A.7. Correlogram Landsat 8 with GLCMs

We assessed the Landsat 8 bands and each of the 8 GLCM texture measures and compared them with each other in groups as recommended (Hall-Beyer, 2017). Ten bands of Landsat 8 were correlated with the 8 GLCMs (Table A.7). The only important information is this Table is in the top 11 rows. Each major column of colors relates to a Landsat band in order from left to right (excluding Band 8) and the very narrow columns relate to a single GLCM measure. There is moderate correlation among GLCM measures in optical bands, but not in the infrared.



	X<0.4		X>0.6<0.8		X=1
	X>0.4<0.6		X>0.8<1.0		

Table A.8. Logistic regression model results in R for cube-root transformed shade coffee and forest texture predictors (Relative Richness and Dominance).

```
> summary(modelesveg3f72)

Call:
glm(formula = veg2 ~ gep3Rt(dBibv) + gep3Rt(dBibvd), family = binomial(link = "logit"),
    data = modeles, maxit = 100)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8162  -0.7337  -0.7292  -0.7234   1.7139

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)    -1.22502    0.08825 -13.880  < 2e-16 ***
gep3Rt(dBibv)   0.23397    0.08959   2.612  0.00901 **
gep3Rt(dBibvd) -0.39739    0.12508  -3.177  0.00149 **
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 66483  on 60943  degrees of freedom
Residual deviance: 66473  on 60941  degrees of freedom
AIC: 66479

Number of Fisher Scoring iterations: 4

==========================
> summary(modelesveg3f66)
```

```
Call:
glm(formula = veg6 ~ gep3Rt(dBibv) + gep3Rt(dBibvd) + gep3Rt(dBibvdi),
    family = binomial(link = "logit"), data = modeles, maxit = 100)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9970  -0.9455  -0.9263   1.4262   1.5777

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)     -1.078535   0.079308 -13.599   <2e-16 ***
gep3Rt(dBibv)    0.136986   0.084353   1.624   0.1044
gep3Rt(dBibvd)   0.212938   0.121580   1.751   0.0799 .
gep3Rt(dBibvdi)  0.003474   0.057890   0.060   0.9521
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 79258  on 60943  degrees of freedom
Residual deviance: 79209  on 60940  degrees of freedom
AIC: 79217

Number of Fisher Scoring iterations: 4
==============================
> summary(modelesveg3f76)

Call:
glm(formula = veg6 ~ gep3Rt(dBibv) + gep3Rt(dBibvd), family = binomial(link = "logit"),
    data = modeles, maxit = 100)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.9978	-0.9453	-0.9260	1.4262	1.5767

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-1.07782	0.07841	-13.745	<2e-16	***
gep3Rt(dBibv)	0.13808	0.08235	1.677	0.0936	.
gep3Rt(dBibvd)	0.21171	0.11982	1.767	0.0772	.

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 79258 on 60943 degrees of freedom

Residual deviance: 79209 on 60941 degrees of freedom

AIC: 79215

Number of Fisher Scoring iterations: 4

Table A.9. Confusion Matrices – Part 1

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_KNN (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	266710	173816	517636	142683	333050	**1433895**	**0.813996**
2	750727	1085608	2626605	333891	699984	**5496815**	**0.802502**
3	380236	61745	1769900	178896	198388	**2589165**	**0.316421**
9	1995223	45499	1634805	1407195	912280	**5995002**	**1**
10	279429	440	73183	318724	77679	**749455**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.927373**	**0.205909**	**0.732729**	**1**	**1**		**0.708033**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000161 (0.807872 - 0.808193)
95% Confidence Interval = +/- 0.000191 (0.807841 - 0.808224)
99% Confidence Interval = +/- 0.000252 (0.807781 - 0.808285)

KAPPA INDEX OF AGREEMENT (KIA)
Using PS_L8_KNN as the reference image ...

Category	KIA
1	-0.051389
2	0.123853
3	0.466266
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	-0.017037
2	0.308974
3	0.128540
4	0.0000
5	0.0000

Overall Kappa = 0.088893

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_MLC (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	14019	7018	12957	12885	11210	**58089**	**0.758663**
2	2342614	1206854	5127664	1274454	1585822	**11537408**	**0.895396**
3	87672	95204	837314	37743	108168	**1166101**	**0.281954**
9	1027011	59166	597786	853073	454021	**2991057**	**1**
10	201105	56	47792	208593	62183	**519729**	**1**
Total	**3672421**	**1368298**	**6623513**	**2386748**	**2221404**	**16272384**	
Error0	**0.996183**	**0.117989**	**0.873585**	**1**	**1**		**0.873517**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000136 (0.873381 - 0.873652)
95% Confidence Interval = +/- 0.000162 (0.873355 - 0.873678)
99% Confidence Interval = +/- 0.000213 (0.873304 - 0.873729)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_MLC as the reference image ...

Category	KIA
1	0.020214
2	0.022400
3	0.524497
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.000248
2	0.594515
3	0.058981
4	0.0000
5	0.0000

Overall Kappa = 0.040520
Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_N_B

(rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	460964	136997	810118	256544	385242	**2049865**	**0.775125**
2	1393917	1158116	4254605	665061	1090167	**8561866**	**0.864736**
3	25275	30389	852516	8523	56898	**973601**	**0.124368**
9	974770	22594	313712	747900	354097	**2413073**	**1**
10	817495	20202	392562	708720	335000	**2273979**	**1**
Total	**3672421**	**1368298**	**6623513**	**2386748**	**2221404**	**16272384**	
Error0	**0.874480**	**0.153608**	**0.871289**	**1**	**1**		**0.758111**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000146 (0.847965 - 0.848257)
95% Confidence Interval = +/- 0.000174 (0.847937 - 0.848285)
99% Confidence Interval = +/- 0.000230 (0.847881 - 0.848341)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_N_B as the reference image ...

Category	KIA
1	-0.001045
2	0.055876
3	0.790259
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	-0.000517
2	0.205823
3	0.073262
4	0.0000
5	0.0000

Overall Kappa = 0.060757
Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_NNET

(rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	1163337	170840	686611	891399	539050	**3451237**	**0.662922**
2	1705789	1009937	3576512	799305	1334895	**8426438**	**0.880147**
3	279231	177603	2231202	91502	201037	**2980575**	**0.251419**
9	523968	8728	127804	599183	146399	**1406082**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.683215**	**0.261260**	**0.663069**	**1**	**1**		**0.729194**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000181 (0.729013 - 0.729375)
95% Confidence Interval = +/- 0.000216 (0.728978 - 0.729410)
99% Confidence Interval = +/- 0.000284 (0.728910 - 0.729478)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_NNET as the reference image ...

Category	KIA
1	0.143744
2	0.039083
3	0.575910
9	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.132759
2	0.306862
3	0.188154
4	0.0000
5	0.0000

Overall Kappa = 0.125587

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_RF (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	93658	57338	119599	64887	71477	**406959**	**0.769859**
2	1349801	1167088	2985855	655190	974449	**7132383**	**0.836368**
3	524512	120136	2790512	261251	563782	**4260193**	**0.344980**
9	1246171	21680	573814	933068	463583	**3238316**	**1**
10	458279	2056	153733	472352	148113	**1234533**	**1**
Total	**3672421**	**1368298**	**6623513**	**2386748**	**2221404**	**16272384**	
Error0	**0.974497**	**0.147051**	**0.578696**	**1**	**1**		**0.711035**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000176 (0.750858 - 0.751211)
95% Confidence Interval = +/- 0.000210 (0.750825 - 0.751245)
99% Confidence Interval = +/- 0.000277 (0.750758 - 0.751311)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_RF as the reference image ...

Category	KIA
1	0.005756
2	0.086848
3	0.418207
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.000507
2	0.728197
3	0.216066
4	0.0000
5	0.0000

Overall Kappa = 0.117400

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_SVM (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	486541	76957	694691	302691	334471	**1895351**	**0.743298**
2	1632289	947045	2515386	885597	960791	**6941108**	**0.863560**
3	403328	326207	2995076	163924	504600	**4393135**	**0.318237**
9	177743	12662	47429	177928	84181	**499943**	**1**
10	972520	5427	370931	856608	337361	**2542847**	**1**
Total	**3672421**	**1368298**	**6623513**	**2386748**	**2221404**	**16272384**	
Error0	**0.867515**	**0.307866**	**0.547812**	**1**	**1**		**0.727842**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000181 (0.727660 - 0.728023)
95% Confidence Interval = +/- 0.000216 (0.727626 - 0.728058)
99% Confidence Interval = +/- 0.000285 (0.727557 - 0.728127)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_SVM as the reference image ...

Category	KIA
1	0.040059
2	0.057159
3	0.463307
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.018119
2	0.463126
3	0.249599
4	0.0000
5	0.0000

Overall Kappa = 0.120916

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_KNN (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	328699	132213	510139	223225	170503	**1364779**	**0.759156**
2	1325087	888433	2560442	707984	775405	**6257351**	**0.858018**
3	666957	286117	3133478	285020	730744	**5102316**	**0.385871**
9	1259076	57636	365126	1087842	477252	**3246932**	**1**
10	92506	2709	52944	77318	67477	**292954**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.910493**	**0.350137**	**0.526817**	**1**	**1**		**0.732506**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000181 (0.732326 - 0.732687)
95% Confidence Interval = +/- 0.000215 (0.732291 - 0.732721)
99% Confidence Interval = +/- 0.000283 (0.732223 - 0.732789)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_KNN as the reference image ...

Category	KIA
1	0.019444
2	0.063243
3	0.349118
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.006107
2	0.430923
3	0.232367
4	0.0000
5	0.0000

Overall Kappa = 0.107772

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_MLC (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	352819	99577	517734	255187	227982	**1453299**	**0.757229**
2	1388489	787104	2624754	638646	874956	**6313949**	**0.875339**
3	251439	384950	2653399	108926	287796	**3686510**	**0.280241**
9	1588634	93727	780052	1302789	763809	**4529011**	**1**
10	90944	1750	46190	75841	66838	**281563**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.903925**	**0.424256**	**0.599313**	**1**	**1**		**0.766771**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000172 (0.766598 - 0.766943)
95% Confidence Interval = +/- 0.000206 (0.766565 - 0.766976)
99% Confidence Interval = +/- 0.000271 (0.766500 - 0.767041)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_MLC as the reference image ...

Category	KIA
1	0.021933
2	0.044332
3	0.527294
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.007380
2	0.306535
3	0.225030
4	0.0000
5	0.0000

Overall Kappa = 0.103094

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_N_B (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	538506	212908	823078	333274	302490	**2210256**	**0.756360**
2	863804	857242	2868873	403932	560862	**5554713**	**0.845673**
3	253239	177744	1848530	104251	298358	**2682122**	**0.310796**
9	1149521	53605	303971	1035969	364836	**2907902**	**1**
10	867255	65609	777677	503963	694835	**2909339**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.853361**	**0.372952**	**0.720856**	**1**	**1**		**0.800528**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000163 (0.800365 - 0.800691)
95% Confidence Interval = +/- 0.000194 (0.800334 - 0.800722)
99% Confidence Interval = +/- 0.000256 (0.800272 - 0.800784)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_N_B as the reference image ...

Category	KIA
1	0.023055
2	0.076720
3	0.475754
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.012433
2	0.433610
3	0.136795
4	0.0000
5	0.0000

Overall Kappa = 0.083503

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_NNET (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	1025607	163696	1131477	671775	662955	**3655510**	**0.719435**
2	1050886	743260	1846694	512083	717725	**4870648**	**0.847400**
3	473710	403592	3317690	175230	475572	**4845794**	**0.315346**
9	1116929	55300	304169	1017278	358535	**2852211**	**1**
10	5193	1260	22099	5023	6594	**40169**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.720720**	**0.456327**	**0.498999**	**1**	**1**		**0.687257**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000189 (0.687068 - 0.687446)
95% Confidence Interval = +/- 0.000225 (0.687032 - 0.687482)
99% Confidence Interval = +/- 0.000297 (0.686960 - 0.687554)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_NNET as the reference image ...

Category	KIA
1	0.070749
2	0.074834
3	0.468078
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.070331
2	0.348600
3	0.289235
4	0.0000
5	0.0000

Overall Kappa = 0.143895

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_RF (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	309377	117093	509335	210750	158067	**1304622**	**0.762861**
2	1381778	955608	3091902	681604	868910	**6979802**	**0.863090**
3	559095	223505	2573829	242229	637130	**4235788**	**0.392361**
9	1344198	69232	403039	1182973	496660	**3496102**	**1**
10	77877	1670	44024	63833	60614	**248018**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.915754**	**0.301000**	**0.611329**	**1**	**1**		**0.763973**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000173 (0.763800 - 0.764147)
95% Confidence Interval = +/- 0.000206 (0.763767 - 0.764180)
99% Confidence Interval = +/- 0.000272 (0.763702 - 0.764245)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_RF as the reference image ...

Category	KIA
1	0.014659
2	0.057705
3	0.338171
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.004383
2	0.472718
3	0.173395
4	0.0000
5	0.0000

Overall Kappa = 0.090268

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_S2_SVM (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	938396	173187	951709	618757	533004	**3215053**	**0.708124**
2	976310	735196	1868805	437248	660178	**4677737**	**0.842831**
3	507611	401788	3430138	200037	555109	**5094683**	**0.326722**
9	1131275	53782	307321	1033344	384269	**2909991**	**1**
10	118733	3155	64156	92003	88821	**366868**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.744468**	**0.462225**	**0.482019**	**1**	**1**		**0.686201**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000189 (0.686012 - 0.686390)
95% Confidence Interval = +/- 0.000226 (0.685976 - 0.686427)
99% Confidence Interval = +/- 0.000297 (0.685904 - 0.686498)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_S2_SVM as the reference image ...

Category	KIA
1	0.085359
2	0.079823
3	0.448890
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	0.072111
2	0.351165
3	0.298123
4	0.0000
5	0.0000

Overall Kappa = 0.146148

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_COSI10 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	68414	81733	118491	45486	50023	**364147**	**0.812125**
2	1851784	1019361	3552395	1026161	1140223	**8589924**	**0.881331**
3	379423	218825	2410902	140231	427162	**3576543**	**0.325913**
9	1078602	46708	417420	866248	485536	**2894514**	**1**
10	294102	481	122921	303263	118437	**839204**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.981370**	**0.254367**	**0.635932**	**1**	**1**		**0.784887**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000168 (0.784719 - 0.785054)
95% Confidence Interval = +/- 0.000200 (0.784687 - 0.785086)
99% Confidence Interval = +/- 0.000263 (0.784624 - 0.785149)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_COSI10 as the reference image ...

Category	KIA
1	-0.048973
2	0.037790
3	0.450255
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	-0.003846
2	0.460922
3	0.184805
4	0.0000
5	0.0000

Overall Kappa = 0.088419

Table A.10. Confusion Matrices – Part 2

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_COSI11 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	64075	84020	107600	42615	50011	**348321**	**0.816046**
2	1830382	1028320	3683663	1006240	1169587	**8718192**	**0.882049**
3	329094	212877	2250274	119911	373415	**3285571**	**0.315104**
9	1089354	41512	436583	820235	494220	**2881904**	**1**
10	359420	379	144009	392388	134148	**1030344**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.982552**	**0.247814**	**0.660189**	**1**	**1**		**0.794479**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000165 (0.794314 - 0.794643)
95% Confidence Interval = +/- 0.000196 (0.794282 - 0.794675)
99% Confidence Interval = +/- 0.000259 (0.794220 - 0.794737)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_COSI11 as the reference image ...

Category	KIA
1	-0.054037
2	0.037006
3	0.468487
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	-0.004055
2	0.465883
3	0.172685
4	0.0000
5	0.0000

Overall Kappa = 0.084552

Error Matrix Analysis of WBVALIDATIONVWR (columns : truth) against PS_L8_COSI11 (rows : mapped)

	1	2	3	4	5	Total	ErrorC
1	64075	84020	107600	42615	50011	**348321**	**0.816046**
2	1830382	1028320	3683663	1006240	1169587	**8718192**	**0.882049**
3	329094	212877	2250274	119911	373415	**3285571**	**0.315104**
9	1089354	41512	436583	820235	494220	**2881904**	**1**
10	359420	379	144009	392388	134148	**1030344**	**1**
Total	**3672325**	**1367108**	**6622129**	**2381389**	**2221381**	**16264332**	
Error0	**0.982552**	**0.247814**	**0.660189**	**1**	**1**		**0.794479**

ErrorO = Errors of Omission (expressed as proportions)
ErrorC = Errors of Commission (expressed as proportions)

90% Confidence Interval = +/- 0.000165 (0.794314 - 0.794643)
95% Confidence Interval = +/- 0.000196 (0.794282 - 0.794675)
99% Confidence Interval = +/- 0.000259 (0.794220 - 0.794737)

KAPPA INDEX OF AGREEMENT (KIA)

Using PS_L8_COSI11 as the reference image ...

Category	KIA
1	-0.054037
2	0.037006
3	0.468487
9	0.0000
10	0.0000

WBVALIDATIONVWR

Category	KIA
1	-0.004055
2	0.465883
3	0.172685
4	0.0000
5	0.0000

Overall Kappa = 0.084552

Table A.11. Landsat 8 Band Correlation Matrix

	B1	B2	B3	B4	B5	B6	B7	B9
B1	1	**0.99**	**0.96**	**0.95**	0.12	**0.82**	**0.87**	-0.2
B2	0.99	1	**0.97**	**0.97**	0.13	**0.84**	**0.9**	-0.17
B3	0.96	0.97	1	**0.97**	0.27	**0.85**	**0.88**	-0.12
B4	0.95	0.97	0.97	1	0.13	**0.9**	**0.93**	-0.11
B5	0.12	0.13	0.27	0.13	1	0.24	0.1	0.04
B6	0.82	0.84	0.85	0.9	0.24	1	**0.96**	-0.11
B7	0.87	0.9	0.88	0.93	0.1	0.96	1	-0.09
B9	-0.2	-0.17	-0.12	-0.11	0.04	-0.11	-0.09	1

Appendix B

R Code for Processing TerrSet Images with Machine Learning Classifiers

The objective was to create code in R that could be run iteratively to classify training set data and testing set data. The following six classifiers were envisioned, and code was prepared for each of them using the raster test data from the TerrSet Image Processing tutorial data set, but have been replaced with some Sentinel 2 work files. The codes are all prepared. One simply must swap out these data for other images and add in additional data frame slots to accommodate additional bands. Some of these codes are obviously part of other's directions, as R-code is setup for everyone to borrow from everyone else to make the new codes work for them. However, these are also unique in that, until this study no one else had complete codes to run TerrSet IDRISI files natively for image classification in R.

MAXIMUM LIKELIHOOD

```
# CLASSIFY Sentinel 2 IMAGES W MAXIMUM LIKELIHOOD IN R
# ========================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_rf_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
# Get the vector of polygon class
```

```
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "mlcfile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeClusner(no_cores))
parLapply(cl, 1:8, rf <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "mlc", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
aveRSTBX(mlc,filename=paste("d:/ps_s2_rf_2348"), overwrite = TRUE)
writeRaster(mlc$map, filenamenb=paste("d:/ps_s2_rf_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END mlc =============
```

K NEAREST NEIGHBOR

```
# CLASSIFY Sentinel 2 IMAGES W K NEAREST NEIGHBOR IN R
# ============================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_nb_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
```

```
# Get the vector of polygon class
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "knnfile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeClusner(no_cores))
parLapply(cl, 1:8, knn <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "knn", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
aveRSTBX(knn,filename=paste("d:/ps_s2_knn_2348"), overwrite = TRUE)
writeRaster(knn$map, filenamenb=paste("d:/ps_s2_knn_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END knn =============
```

NEURAL NETWORK

```
# CLASSIFY Sentinel 2 IMAGES W NEURAL NETWORK IN R
# ========================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_nnet_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
# Get the vector of polygon class
```

```
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "nnetfile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeCluster(no_cores))
parLapply(cl, 1:8, nnet <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "nnet", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
saveRSTBX(nnet,filename=paste("d:/ps_s2_nnet_2348"), overwrite = TRUE)
writeRaster(nnet$map, filename=paste("d:/ps_s2_nnet_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END nnet =============
```

SUPPORT VECTOR MACHINES

```
# CLASSIFY Sentinel 2 IMAGES W SUPPORT VECTOR MACHINES IN R
# ========================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_svm_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
# Get the vector of polygon class
```

```
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "svmfile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeCluster(no_cores))
parLapply(cl, 1:8, svm <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "svmLinear", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
saveRSTBX(svm,filename=paste("d:/ps_s2_svm_2348"), overwrite = TRUE)
writeRaster(svm$map, filename=paste("d:/ps_s2_svm_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END svm =============
```

RANDOM FORESTS

```
# CLASSIFY Sentinel 2 IMAGES W RANDOM FORESTS IN R
# ========================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_rf_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
# Get the vector of polygon class
```

```
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "rffile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeClusner(no_cores))
parLapply(cl, 1:8, rf <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "rf", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
aveRSTBX(rf,filename=paste("d:/ps_s2_rf_2348"), overwrite = TRUE)
writeRaster(rf$map, filenamenb=paste("d:/ps_s2_rf_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END rf =============
```

NAÏVE BAYES

```
# CLASSIFY Sentinel 2 IMAGES W NAÏVE BAYES IN R
# ============================================================
# install.packages("e1071", dependencies = TRUE)
# install.packages("raster", dependencies = TRUE)
# install.packages("caret", dependencies = TRUE)
# install.packages("RStoolbox", dependencies = TRUE)
# install.packages("ddalpha", dependencies = TRUE)
# install.packages("snow"), dependencies = TRUE)
library(RStoolbox)
library(e1071)
library(raster)
library(ddalpha)
library(caret)
library(snow)
library(parallel)
# Variables
shapePath <- 'd:/data3/s2/wb_tsitesu.shp'
brickPath <- 'd:/data3/s2'
lcPath <- 'd:/data3/s2/wb_s2_nb_pred'
setwd(brickPath)
getwd()
# Load objects
wbapbb_b2 <- raster("wbapbb_b2.rst")
wbapbb_b3 <- raster("wbapbb_b3.rst")
wbapbb_b4 <- raster("wbapbb_b4.rst")
wbapbb_b8 <- raster("wbapbb_b8.rst")

brick <- brick(stack(wbapbb_b2, wbapbb_b3, wbapbb_b4, wbapbb_b8))
imgp <- brick
# Get the vector of polygon class
```

```
trainData <- shapefile(shapePath)
responseCol <- "class"
lc = "nbfile"

# Train the classifier

# Calculate the number of cores
no_cores <- detectCores() - 1
no_cores
# Initiate cluster
system.time(cl <- makeCluster(no_cores))
parLapply(cl, 1:8, nnet <- superClass(imgp, trainData, valData = NULL, # RStoolbox
                    responseCol = "class", nSamples = 1000,
                    polygonBasedCV = FALSE, trainPartition = NULL,
                    model = "nb", tuneLength = 3, kfold = 5, minDist = 2,
                    mode = "classification", predict = TRUE,
                    predType = "raw", filename = lc, verbose = TRUE,
                    overwrite = TRUE), arg = NULL)
saveRSTBX(nb,filename=paste("d:/ps_s2_nb_2348"), overwrite = TRUE)
writeRaster(nb$map, filenamenb=paste("d:/ps_s2_nb_2348"), format='RST',
datatype='INT1U', overwrite=TRUE)
# ============ END nb =======
```

Appendix C

Publisher's Authorization

Taylor & Francis
Taylor & Francis Group

Our Ref: LA/TLUS/P17/1698

12 November 2017

Dear Peter Schlesinger,

Material requested: Peter Schlesinger, Carlos L. Muñoz Brenes, Kelly W. Jones & Lee A. Vierling (2017) The Trifinio Region: a case study of transboundary forest change in Central America, Journal of Land Use Science, 12:1, 36-54

Thank you for your correspondence requesting permission to reproduce the above mentioned material from our Journal in your printed thesis entitled "Understanding Pattern Land-cover and land-use change in the Trifinio Region of Central America" and to be posted in the university's repository - University of Idaho.

We will be pleased to grant permission on the sole condition that you acknowledge the original source of publication and insert a reference to the article on the Journals website: http://www.tandfonline.com

This is the authors accepted manuscript of an article published as the version of record in Journal of Land Use Science © 09 Dec 2016 http://tandfonline.com/doi/full/10.1080/1747423X.2016.1261948

This permission does not cover any third party copyrighted work which may appear in the material requested.

Please note that this license does not allow you to post our content on any third party websites or repositories.

Thank you for your interest in our Journal.

Yours sincerely

Lee-Ann

Lee-Ann Anderson – Senior Permissions & Licensing Executive, Journals
Taylor & Francis Group
3 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN, UK.
Tel: +44 (0)20 7017 7932
Fax:+44 (0)20 7017 6336
Web: www.tandfonline.com
e-mail: lee-ann.anderson@tandf.co.uk

Taylor & Francis Group
an informa business

Taylor & Francis is a trading name of Informa UK Limited,
registered in England under no. 1072954