Modelling PES (Payments for Ecosystem Services) to prevent deforestation in tropical montane cloud forests (TMCFs) and meet the socioeconomic needs of their farmers in Latin America

Pablo Martin

THÈSE

Pour obtenir le grade de

Docteur

Délivré par **l'Université Paul-Valéry Montpellier 3** et **Universidad Politécnica de Madrid**

Préparée au sein de l'école doctorale ED 60
Et de l'unité de recherche INNOVATION

Spécialité : **Sociologie**

Présentée par **Pablo MARTIN**

Modelling PES (Payments for ecosystem Services) to prevent deforestation in tropical montane cloud forests (TMCFs) and meet the socioeconomic needs of their farmers in Latin America

Soutenue le 26 novembre 2021 devant le jury composé de

Mme Susana MARTIN FERNANDEZ, Associate Professor, Universidad Politécnica de Madrid	Présidente du jury
Mme Stéphanie CARRIERE, Directrice de recherche, Université Montpellier 3	Rapportrice
M. Jésus Javier LITAGO LAVILLA, Professeur, Universidad Politécnica de Madrid	Examinateur
Mme Claire BERNARD-MONGIN, Chargée de recherche, Université de Montpellier	Examinatrice
M. Jean-François LE COQ, Chargé de recherche, CIRAD	Examinateur
M. Fernando MONTES PITA, Chargé de recherche, CSIC-INIA	Examinateur
Mme Marie RIVERA MENDEZ, Chargée de recherche, Université d'Evora	Examinatrice

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería
Agronómica, Alimentaria y de Biosistemas

Modelling PES (Payments for Ecosystem Services) to prevent deforestation in tropical montane cloud forests (TMCFs) and meet the socioeconomic needs of their farmers in Latin America

DOCTORAL THESIS

Pablo Martín Ortega

Technical Forest Engineer (UVA)
MSc European Forestry (UEF)

2021

UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA, ALIMENTARIA Y DE BIOSISTEMAS

DEPARTAMENTO DE INGENIERÍA Y GESTIÓN FORESTAL Y AMBIENTAL

Modelling PES (Payments for Ecosystem Services) to prevent deforestation in tropical montane cloud forests (TMCFs) and meet the socioeconomic needs of their farmers in Latin America

DOCTORAL THESIS

Pablo Martín Ortega
Technical Forest Engineer (UVA)
MSc European Forestry (UEF)

Thesis Co-Supervisors: Luis Gonzaga García-Montero and Nicole Sibelet

Madrid, 2021

Tribunal nombrado por el Sr. Rector Magfco. de la Universidad Politécnica de Madrid, el día.....de.....de 20....

Presidente: _____

Vocal: _____

Vocal: _____

Vocal: _____

Secretario: _____

Suplente: _____

Suplente: _____

Realizado el acto de defensa y lectura de la Tesis el día.....de.....de 20 ...
en la E.T.S.I. / Facultad.....

Calificación .....

EL PRESIDENTE LOS VOCALES

EL SECRETARIO

Contents

i. Preface

This dissertation is written in partial fulfillment of the requirements for obtaining a degree of Doctor of Philosophy from the Department of Forest and Environmental Engineering and Management, Technical University of Madrid, and the University Paul-Valéry Montpellier 3. The original research presented in this dissertation was carried out between September 2015, and September 2021. The work was financially supported by the Agricultural Transformation by Innovation (AgTraIn) program and the University of Copenhagen. Broadly, the thesis focuses on the analysis of remote sensing tools to monitor vegetation and predict deforestation risk, but also uses a socioeconomic approach that supports and identifies the drivers of the predicted deforestation. The work has culminated in one paper with the following title:

Martín-Ortega, P., García-Montero, L.G., Sibelet, N., 2020. Temporal patterns in illumination conditions and its effect on vegetation indices using Landsat on Google Earth Engine. Remote Sens. 12. https://doi.org/10.3390/rs12020211

ii. Acknowledgments

First of all, I would like to thank my two supervisors Luis Gonzaga García-Montero and Nicole Sibelet for their support not only academically but in many other personal and life aspects. I would also like to thank the Agricultural Transformation by Innovation (AgTraIn) program and the European Commission, which provided the funding to make this research possible. Many friends, researchers, and professors helped me in many ways with their ideas across the different institutions AgTraIn, UPM, SupAgro, CIRAD, and CATIE, it would be a long list, but thanks to all of you that listened to me, gave me any feedback, shared a coffee and made constructive comments about my research. A Ph.D. involves a lot of time lost in paperwork, thanks to all the administrative staff of these institutions who guided and helped me, with special thanks to Myriam Perez-Dumoulin for her support.

My most sincere and emotional thanks to all the people I met in Costa Rica, where I had wonderful experiences and learned so much. Thanks to the farmers for welcoming me into their homes, meeting their families, and sharing delicious coffee with me, because I understood my research thanks to them. Learning from Costa Rica people and its landscapes has been a unique life experience that I will always remember, the real inspiration for my research has come from this beautiful country. Special thanks to CATIE, the administrative staff that helped me and welcomed me so nicely, and all the people and friends I met there, including the students that helped me with the interviews. CATIE is a wonderful institution in many ways, it was a real privilege to be there.

Finally thanks to my family, friends and special thanks to Eva for being by my side and being patient until I finally closed this stage.

List of figures

List of tables

iii. Abstract

The development of reliable tools to predict deforestation risk is key to demonstrate efficiency and additionality when targeting protected areas or designing programs of payments for environmental services (PES). It is also important to have a multidisciplinary approach that combines biophysical and socioeconomic sciences when analyzing the effect of environmental policies. In this research, a predictive model of deforestation risk using machine learning techniques was developed for a study site in Costa Rica based on the analysis of historical deforestation patterns throughout the period 2000-2018. Historical and predicted deforestation patterns were analyzed within protected areas and farms participating in the national PES program. The interpretation of the predictive model and drivers of deforestation was completed with socioeconomic information collected through semi-structured interviews from 67 farmers participating and not participating in the PES program. Finally, a methodological analysis of the remote sensing techniques employed to monitor vegetation revealed how topography is an important factor that may have implications when it comes to monitoring forest cover changes throughout time.

The historical vegetation loss rate in the area was low (-0.14% y^{-1}) compared with countries with the highest deforestation rates in the same period. Besides, most of the vegetation loss observed occurred outside protected areas. Regardless of whether the farms participated or not in the PES program, historical vegetation loss rates were very low as well. 94% of their forest remained undisturbed throughout the period 2000-2018. In general, deforestation was higher for farms not participating in the PES program and occurred due to small forest adjustments around productive areas, instead of land cover change due to extensive agricultural transformation. The results indicate that the low deforestation rates found in farms participating in the PES program could be explained because forest lands would be spatially biased towards lower pressure, where opportunity cost is low.

The predictive model showed an accuracy of 0.89 in predicting vegetation loss in the study site. Additionally, it provides biophysical and spatially explicit information to understand the drivers of forest loss, and the locations where this is likely to occur, which can improve decisions taken when designing environmental policies. Topography and accessibility were the main factors influencing deforestation in the area due to the mountainous nature and irregular terrain where cloud forests are located. In general, forests are located far from main

roads in complex terrain and they are under low threat of deforestation. Protected areas and farms that received PES are generally located in remote areas where the model predicts a low risk of deforestation and this should be used to question the efficiency and additionality of focusing economic resources on these areas. The interviews showed that the opportunity cost of transforming forests into a more profitable land use might be very low in remote areas and emergent forest uses such as ecotourism might also be preventing forest clearing.

Finally, it was also found an important effect of the topography on the satellite sensors employed to monitor vegetation in mountainous areas, which can bias the estimates of important environmental services derived from the vegetation throughout time. A novel approach was developed to use Landsat temporal series to evaluate changes in the illumination conditions and vegetation indices in forested areas in irregular terrain.

iv. Resumen

El desarrollo de herramientas fiables para predecir el riesgo de deforestación es clave para demostrar la eficiencia y adicionalidad al diseñar áreas protegidas o programas de pagos por servicios ambientales (PSA). También es importante contar con un enfoque multidisciplinar que combine las ciencias biofísicas y socioeconómicas a la hora de analizar el efecto de las políticas ambientales. En esta investigación, se desarrolló un modelo predictivo del riesgo de deforestación mediante técnicas de aprendizaje automático para una zona de estudio en Costa Rica basado en el análisis de patrones históricos de deforestación durante el período 2000-2018. Se analizaron los patrones de deforestación históricos y los pronosticados dentro de áreas protegidas y fincas que participaron en el programa nacional de PSA. La interpretación del modelo predictivo y los factores que impulsan la deforestación se completaron con información socioeconómica recopilada a través de entrevistas semiestructuradas a 67 agricultores participantes y no participantes en el programa de PSA. Por último, un análisis metodológico de las técnicas de teledetección empleadas para monitorizar la vegetación reveló cómo la topografía es un factor importante que puede tener implicaciones cuando se trata de monitorizar los cambios de la cubierta forestal a lo largo del tiempo.

La tasa histórica de pérdida de vegetación en el área fue baja (-0.14% y^{-1}) en comparación con los países con las tasas de deforestación más altas en el mismo período. Además, la mayor parte de la pérdida de vegetación observada ocurrió fuera de las áreas protegidas. Independientemente de si las fincas participaron o no en el programa de PSA, sus tasas históricas de pérdida de vegetación también fueron muy bajas. El 94% de su bosque permaneció inalterado durante el período 2000-2018. En general, la deforestación fue mayor para las fincas que no participaban en el programa de PSA y se produjo debido a pequeños ajustes de la cubierta forestal alrededor de los terrenos productivos, en lugar de cambios en la cobertura de la tierra debido a una extensa transformación agrícola. Los resultados indican que las bajas tasas de deforestación encontradas en las fincas que participan en el programa de PSA podrían explicarse porque las tierras forestales tienden a ubicarse en areas con menor presión, donde el costo de oportunidad es bajo.

El modelo predictivo mostró una precisión de 0,89 en la predicción de la pérdida de vegetación en el sitio de estudio. Además, proporciona información biofísica y espacialmente

explícita para entender los factores que explican la pérdida de bosque y las zonas donde es probable que esto ocurra. Esto es fundamental para mejorar la toma de decisiones a la hora de diseñar políticas medioambientales. La topografía y la accesibilidad fueron los principales factores que influyeron en la deforestación en la zona debido a la naturaleza montañosa y el terreno irregular donde se ubican los bosques nubosos. En general, los bosques están ubicados lejos de las carreteras principales en terrenos complicados y están bajo una baja amenaza de deforestación. Las áreas protegidas y las fincas que recibieron PSA generalmente están ubicadas en áreas remotas donde el modelo predice un bajo riesgo de deforestación y esto debe usarse para cuestionar la eficiencia y adicionalidad de focalizar recursos económicos en estas áreas. Las entrevistas mostraron que el costo de oportunidad de transformar los bosques en un uso de la tierra más rentable podría ser muy bajo en áreas remotas y los usos forestales emergentes, como el ecoturismo, también podrían estar impidiendo la pérdida de bosques.

Por último, se encontró un efecto importante de la topografía sobre los sensores satelitales empleados para monitorizar la vegetación en áreas montañosas, lo que puede sesgar las estimaciones de importantes servicios ambientales derivados de la vegetación a lo largo del tiempo. Se desarrolló un enfoque novedoso para utilizar series temporales de Landsat para evaluar los cambios en las condiciones de iluminación y los índices de vegetación en áreas boscosas en terreno irregular.

1. Introduction

1.1. Context, challenges, and stakes

Greenhouse gases released by deforestation have been identified as one of the main human causes of climate change (IPCC, 2021). Deforestation continues and the percentage of the terrestrial forest area decreased from 32.5% to 30.8% in the three decades between 1990 and 2020, although the average rate of net deforestation has decreased by 40% comparing the decades 1990–2000 and 2010–2020. Even though deforestation rates show a decreasing trend, forests continue to be important not only for their contribution to greenhouse gases but also for the diversity of their ecosystems, species, and genetic material that underpin life on Earth (FAO and UNEP, 2020).

Efforts to conserve forests and other ecosystems started a long time ago using the figure of protected areas (PA). The origins of “modern” PA are to be found in the nineteenth century, with significant growth in number and extent throughout the twentieth century (Phillips, 2004). Initially, PA were established to be a representative set of the ecosystems of the world and remained relatively isolated with strong protection measures. Today a PA must coexist with the impact of human activities, including their synergies (Lovejoy, 2006).
Recent studies indicate that due to this protection, deforestation in the world´s PA has decreased but not disappeared when it was compared with control areas outside PA with similar characteristics. Moreover, only 6.5%—rather than 15.7%—of the world’s forests turned out to be effectively protected (Wolf et al., 2021). It has been also criticized that global conservation goals focus solely on PA area expansion, yet much of this expansion has been inadequately targeted (Barnes et al., 2018).

A more recent and widely used figure in the field of conservation and forest protection are ecosystem services and the market-based instruments known as Payment for Ecosystem Services (PES) Schemes. The origins of ecosystem services are to be found in the 1970s, in order to increase public interest in beneficial ecosystem functions. Ecosystem services have evolved to a commodification process that has been finally refined with the creation of institutional structures and schemes (public or private) that allow the monetary transactions and exchanges of ecosystem services (Gómez-Baggethun et al., 2010).

PES were proposed trying to replace other tools such as PA that were not having the desired effect on conservation (Ferraro and Kiss, 2012). However, the impact of PES continues to be questioned from a variety of perspectives and cannot be considered as the most cost-effective policy option to achieve environmental goals (Muradian et al., 2013). From the deforestation perspective, many studies have focused on the effect of PES in reducing deforestation rates and discover that, after reviewing several PES programs in different countries, their impact on reducing deforestation was low (Börner et al., 2017).

It is important that PA and PES, which are designed as policy environmental tools, can be properly evaluated and respond to the principles of effectiveness, efficiency, and additionality.

Effectiveness can be defined as: “The increase in ecosystem service provision or conservation effect against the absence of PES or PA that is achieved for a given budget” (Börner et al., 2017).

Efficiency can be defined as: “The maximization of an ecosystem service provision or conservation effect that is achieved with the minimum given budget, and maximizing as well the participation of those actors that constitute a credible threat or actively increase these provisions or effects” (Wunder, 2007).

Additionality can be defined as: “The incremental ecosystem service provision or conservation effect vis-à-vis predefined baselines” (Wunder, 2007).

Empirical designs are critical to evaluate and monitor environmental policies and provide the evidence needed to properly evaluate the principles above mentioned (Ferraro, 2009). Moreover, these designs must incorporate biophysical and socioeconomic approaches since it is a combination of these 2 factors which are responsible for land management and land-use decisions that ultimately affect deforestation and forest conservation (Karsenty and Ezzine-de-blas, 2016).

The use of remote sensing has been identified as one of the most cost-effective tools to monitor deforestation and its main drivers in the last 20 years in developing countries (Leblois et al., 2017). Additionally, remote sensing has been used to spatially map a

multitude of ecosystem services with the advantage to perform spatially continuous and frequent observations, which is key for monitoring purposes (De Araujo Barbosa et al., 2015). Spatial analysis using forest cover as a proxy has been widely employed to evaluate the effectiveness of PA or PES policies (Andam et al., 2008; Havinga et al., 2020; Mokondoko et al., 2018; Pfaff et al., 2015, 2009a, 2008; Robalino and Pfaff, 2013; Sanchez-Azofeifa et al., 2003). One of the most interesting results of these evaluations is the level of threat faced by PA or targeted forests receiving PES based on their location, and how this greatly affects the effectiveness, efficiency, and additionality of these policies. Most of these results were obtained after the PA or PES were enforced, but little research has been done predicting where deforestation is likely to occur, which could help in the design phase, previous to the enforcement of the policy or program.

In order to generate accurate predictions and deforestation models, machine learning (ML) algorithms are presented as an interesting solution. These algorithms can handle large amounts of data, multiple variables, and non-linear or complex relations between biophysical and socioeconomic variables without the need for any transformation (Mayfield et al., 2017). Models to predict deforestation have been successfully implemented in areas such as Borneo (Cushman et al., 2017), India (Bera et al., 2020), Mexico, Madagascar (Mayfield et al., 2017), or South Africa (Dlamini, 2016). One important advantage of these methods is that they can provide spatially continuous and explicit deforestation risk maps.

But deforestation cannot be disconnected from socioeconomic drivers, such as economic development, agricultural activity, or population pressure (Leblois et al., 2017). Understand forest owners' perceptions, decisions and needs have been identified as a key factor for a better design of environmental policies. Understand factors such as opportunity cost, wealth, level of education, transaction costs, alternatives to the use of forest or even the voluntary participation nature of most of the PES programs, is essential when the environmental impact has to be evaluated (Allen and Colson, 2019; Fiorini et al., 2020; Pfaff et al., 2008 ; Salzman et al., 2018; Sheng et al., 2019; Zbinden and Lee, 2005)

Finally, the use of remote sensing techniques to characterize and monitor forests and the ecosystem services derived from them has greatly increased in the last years (De Araujo Barbosa et al., 2015). These techniques can evaluate the state and health of the vegetation by using relatively simple indexes, such as the normalized difference vegetation index

(NDVI), and then infer from these indexes ES such as carbon sequestration, water yield, or biodiversity (Cao et al., 2016; Cord et al., 2017). However, the characteristics of the sensors employed need to be fully understood and calibrated to interpret the results. Temporal and spatial resolutions of satellite sensors, as well as the effect of sun-sensor geometry, topography, or seasonality, affect the surface reflectance coming from the vegetation (Nagol et al., 2014; Van Den Hoek et al., 2021). Overlooking this can bias the interpretation of deforestation or the ES provided by forests.

1.2. Forest dynamics and conservation in Costa Rica

Costa Rica has gained a net area of 128,000 hectares of forest (4.40%) during the period 1990-2020 (FAO, 2020). However, the country experienced yearly deforestation rates of -2.90% between the years 1981-1990 (FAO, 1990). During the years 1950-1960, Government policies promoted the conversion of forests to cropland or pasture, because of the growing population and the increase in food demand (Sánchez-Azofeifa et al., 2003). Although agricultural production has increased continuously since 1965, technology improved, decreasing the land area needed for crops. After the collapse of both meat exports and the coffee crisis experienced in the 1980s, reforestation occurred at the expense of pasture lands and traditional crops (Jadin et al., 2016). While some authors have pointed out that the forest in Costa Rica is now on a period of recovery and even predicted a continuous forest expansion (Jadin et al., 2016; Stan and Sanchez-Azofeifa, 2019; Vallet et al., 2016), the worrying deforestation rates that the country experienced in the past, triggered mechanisms and environmental policies for forest conservation.

1.2.1. Protected wildlife areas and deforestation

Since the 1960s, Costa Rica has designated more than 150 protected areas (Andam et al., 2008). Between 1974 to 1978, the creation of Protected Wildlife Areas (PWA) expanded from 3 to 12% of the area of the country, partly, trying to offset deforestation trends, today, the percentage is 25% (Robalino et al., 2017). The biodiversity law 7788 from 1998 defines PWA as: ”A defined geographic area, officially declared and designed with a management category based on its natural, cultural and socio-economic importance to accomplish specific management and conservation objectives”. It states that the PWA are selected based on their ability to protect biodiversity, soil, hydrological resources, and other environmental

services. Among all PWA in the country, one of the most important in terms of biodiversity, uniqueness, and extent is the Tapantí National park, the best-preserved cloud forest in the country (Bernard et al., 2009).

However, the establishment and location of the PWAs have a key role in the effectiveness of achieving the pursued decrease in deforestation rates (Andam et al., 2008; Pfaff et al., 2009a, 2009b; Sanchez-Azofeifa et al., 2003). To effectively allocate scarce economic resources for conservation, it cannot be assumed that protected areas simply will decrease deforestation (Pfaff et al., 2009b). Most of the evaluations measuring the effect of conservation rely on indirect estimates of deforestation by just comparing protected and unprotected areas, but frequently the characteristics of protected areas are very specific, and not considering this would lead to biased analysis (Andam et al., 2008). Globally, protected areas are biased towards higher elevations, slopes, and distances to roads and cities, facing in general low land conversion pressures (Joppa and Pfaff, 2010).

Effectiveness and additionality, this is, that the protected area is avoiding deforestation versus the absence of the protection, should guide decisions to target conservation areas (Andam et al., 2008; Wunder, 2007). In Costa Rica, previous studies have shown that the additionality provided by the protected areas during the period 1960-1997 has been low compared with unprotected areas of similar characteristics (Andam et al., 2008; Pfaff et al., 2009a, 2009b). Factors behind the low deforestation threat are that the location of protected areas in Costa Rica is biased towards remote areas, at high altitudes and with poor soils, where land-use change for other uses than having forest would not be very attractive (Joppa and Pfaff, 2010; Sanchez-Azofeifa et al., 2003). In Costa Rica, it was found that PWAs had a greater effect on avoiding deforestation when they were located close to main urban areas and roads and on lower slopes (Pfaff et al., 2009a; Pfaff et al., 2014, 2015). Location of protected areas is frequent in less productive lands and further from the forest frontier (Andam et al., 2008).

Empirical analysis on deforestation in Costa Rica accounting for the bias on the location of PWAs by using matching techniques showed that only 7-9% of the protected forest would have been deforested between 1960 and 1996 in the absence of protection. Previous studies would have overestimated deforestation rates by a factor of three or more simply because they would not be considering that the location of the protected areas is biased (Andam et

al., 2008). Later studies between 1986-1997 decreased these estimates to 1-2% (Pfaff et al., 2009a). Conservation areas showed negligible deforestation rates inside them in a more extended analysis between 1960 and 1997 but increased rates in adjacent buffer areas of 10 km (Sánchez-Azofeifa et al., 2003). Other studies between 1960 and 1997, estimated that only 13.5% of previously unforested lands inside protected areas were reforested because they were protected (Andam et al., 2013). Recent global estimates showed that only 6.5% of the world's forests are protected when compared with unprotected forests of similar characteristics, and this is quite far from the Aichi Convention on Biological Diversity's 2020 Target of 17% (Wolf et al., 2021).

1.2.2. Program of payment for environmental services and deforestation

The first experiences related to the promotion and conservation of forests in Costa Rica go back to 1969, incentivizing or compensating the payment of taxes of private owners in exchange for reforesting land. With the enactment of Forest Law 7575 in 1996, the concept of payment for environmental services (PES) was created and also the National Forest Financing Fund (FONAFIFO), which is the organism in charge of the administrative procedures of implementing, monitoring, selecting beneficiaries and delivering the payments (Sánchez and Navarrete, 2017). The 7575 Law also included a ban on forest clearing and the PES program tried to offset that (Daniels et al., 2010; Morse et al., 2009). The PES program provides payments in recognition of the environmental services that private forests provide to society (Arriagada et al., 2009). Environmental services recognized in the Forest Law 7575 for payment are defined as:

"Those that forests and forest plantations provide and that directly affect the protection and improvement of the environment".

Specifically, these services include a) Mitigation of greenhouse gas emissions; b) the protection of water for urban, rural, or hydroelectric use; c) the protection of biodiversity to conserve and use it in a scientific, sustainable, and pharmaceutical way, and d) scenic beauty (MINAE, 1996).

The PES program offers different payment modalities focused on reforestation, forest management, agroforestry, or forest protection. Between 1997 and 2016, 90% of the area

receiving payments (1,050,135 hectares) was under the forest protection modality (Porras and Asquith, 2018). Forest protection modality has been widely studied for its importance and because its effect can be analyzed by monitoring forest cover changes using remote sensing or geographical information systems (GIS). As a private owner, participation in the PES program is voluntary, but FONAFIFO periodically establishes a valuation matrix using environmental and socioeconomic features to give a score to each farm willing to participate in the forest modality (Table 1).

Table 1. FONAFIFO valuation matrix to establish priority criteria in the selection of beneficiary farms of the PES program (source: Executive Decree Nº 39871, La Gaceta, 2016)

Criteria Number	Priority criteria	Score
1	Forests in private farms located inside Protected Wildlife Areas	115
	Forests inside indigenous territories in the country	
2	Forests on farms located in defined areas within sites of conservation importance	110
	Forests on farms located within the officially established Biological Corridors	
3	Forests that protect sources for water supply, mainly for the population's consumption (based on information provided by Acueductos y Alcantarillados, Asadas, or with a note from municipalities that administer aqueducts)	105
4	Forests outside of any of the above priorities	55
I	Forests for protection that comply with the provisions of the previous points, and that have signed contracts for the payment of environmental services for forest protection in previous years, contracts that conclude their period of validity will also be considered for these purposes.	10 additional
II	Forests on farms located in districts with a Social Development Index (IDS) less than 43.4% according to the determination made by MIDEPLAN (2013).	10 additional
III	Forests in any of the previous priorities, with an application to enter the Payment for Environmental Services Program in areas smaller than 50 hectares. These points only apply if the area of the farm is equal to or less than 50 hectares	25 additional
IV	Forests in any of the priorities a, b, c, d, e, and f, with an application to enter the Payment for Environmental Services Program that have less than 100 hectares of the real folio and a proposed PES area of 50 hectares maximum, for projects of Forest Protection processed by organizations with a current agreement with the National Forest Financing Fund, which are not included in the previous paragraph	10 additional

Because the matrix is the result of overlaying different spatial layers that contain the information defining the priority criteria (FONAFIFO, personal communication, 2017), farms finally accessing the PES program will be chosen based on a strong spatial component. Proximity to protected areas, biological corridors, or areas of conservation importance are key factors to get the highest score since forests in these areas have a high natural and socio-economic importance. While it is not in question that it is important to protect the forest in these areas, an additional question would be whether these forests are facing any deforestation threat that would affect the provision of environmental services. As previously shown for the establishment of PWAs, additionality and deforestation were determined to be

low (Andam et al., 2008; Pfaff et al., 2009a, 2009b), therefore, it is legitimate to question the efficiency and additionality of the PES program if it is targeting low threatened locations. Although additionality and efficiency are not explicitly part of the PES program design, any PES program should aim to maximize these two components (Daniels et al., 2010; Robalino and Pfaff, 2013; Wunder, 2007).

Evaluation of the efficiency and additionality of forest conservation in the PES program involves two dimensions: biophysical and socio-economic. Biophysical factors are characteristics that determine the presence of the forest in a given area, such as soil, topography, or climate. The socio-economic dimension involves decisions that drive governments or individuals, through policies or personal decisions to convert forest land on other land use, for economic factors, national policies cultural or social factors (Geist and Lambin, 2002; Jadin et al., 2016; Meyfroidt et al., 2010).

The impact of the PES program on forest cover has been evaluated between participants and non-participants using matching techniques to compare farms between both groups with similar characteristics. Identifying the effect of the program would be straightforward by just comparing farms participating and non-participating, however, as it was shown before with the valuation matrix (Table 1), there are some objectives and restrictions that make that farms receiving PES are not randomly located (Robalino et al., 2011). Estimates between 1997-2000 found that, in comparison with farms not enrolled in the program, only 0.20% of the forest land enrolled in the PES program would have been annually deforested in the absence of the payment. When matching techniques were used to compare groups of untreated non-PES locations more ‘similar’ to the PES locations, the estimates decreased to a 0.08% annual rate, which is very low (Pfaff et al., 2008). A similar comparison in different regions in Costa Rica between the years 2000-2005 found that the PSA program decreased deforestation by 0.61 to 0.69% annually, but avoiding deforestation only on 3-3.5% of the farms enrolled which is quite low. The results also showed that deforestation rates were unevenly distributed across the different regions studied, which suggests the importance of targeting deforestation threat, and local factors affecting its variation (Robalino et al., 2011). Other studies also found that forest cover would be similar in PES and non-PES farms in the absence of the program (Sierra and Russman, 2006).

Analysis has demonstrated that farms participating in the program were located in areas facing low deforestation threats, far from urban areas, and national roads (Robalino and Pfaff, 2013). Frequently farms enrolled in the PES program were located in more remote areas, where opportunity costs are lower (Robalino et al., 2011) and are larger than non-PES farms, which is positively associated with a higher proportion of forest land within the farm (Sierra and Russman, 2006; Zbinden and Lee, 2005). In tropical forests in the Amazon region, location of farms far from roads, with steep slopes, or with a high proportion of forest are factors that negatively affect land value (Chomitz et al., 2005; Holland et al., 2016; Merry et al., 2008; Sills and Caviglia-Harris, 2009). Biophysical factors, therefore, influence the economic value of a farm and influence the decisions of its owners.

Socio-economic analysis based on interviews with PES program participants and non-participants in Costa Rica helped to understand the preferences of owners to participate or not in the program. In general, low opportunity costs in land use and a higher proportion of land occupied by forest were positively associated with participation in the program (Arriagada et al., 2009). Additional factors such as human capital, access to information about the PES program, and the possibility to afford the transaction costs associated with the program participation were also positively correlated with participation (Morse et al., 2009; Zbinden and Lee, 2005). Although the landscape imposes natural conditions and limitations to the activities that can be developed on particular land use, which influences economic decisions, the analysis of the social dimension allows us to understand ecological, social, and cultural values that affect the decisions of the owners to enroll in environmental programs or that influence their attitude towards forest conservation (Sibelet et al., 2017).

1.3. Predictive models to evaluate forest risk loss

Worldwide and in Costa Rica, most of the evaluations targeting the efficiency and additionality of forest conservation policies on forest cover, such as the establishment of PWAs or the PES program have been made using matching techniques (Pfaff et al., 2009, 2014; Robalino et al., 2011; Robalino and Pfaff, 2013; Sierra and Russman, 2006; Wolf et al., 2021). Although these techniques offer reliable results, still lack the incorporation of other factors that can bias the interpretation of the results. They do not incorporate for example if the non-PES farms were eligible or not to participate, or social aspects such as access to information about the program or the effect of tourism as a driver of forest conservation

(Daniels et al., 2010). Matching techniques, compared similar groups of farms or PWAs with areas of similar characteristics, but that sometimes were far in location (i.e., across different provinces), lacking important information that can affect socio-economic factors that influence deforestation rates at the local level (Pfaff et al., 2008; Robalino et al., 2011; Wolf et al., 2021). In addition, most of these studies have employed historical deforestation rates, or focused on a short period of time, and projected them forward in a linear fashion, which may introduce biases on predictions, since deforestation rates may respond to complex non-linear variables in time and space (Daniels et al., 2010).

A thorough analysis of deforestation trends and future projections in Costa Rica using biophysical and socioeconomic variables predicted a continuous increase in forest cover, even in the most adverse scenarios, this, at least, demands a better approach of environmental policies for the protection of forests (Stan and Sanchez-Azofeifa, 2019). Analysis like this highlights the importance of identifying areas at higher risk of future deforestation according to patterns observed in historical deforestation trends. In this sense, targeting areas to avoid deforestation will be much improved (Aguilar-Amuchastegui et al., 2014)

The prediction of deforestation risk has been carried out through the use of simple ordinary least square regressions (OLS) in the humid tropics and the Amazon basin, to more complex models such as binary logistic regression in India (Bera et al., 2020), or maximum entropy (MaxEnt) in the Peru Amazon forests (Aguilar-Amuchastegui et al., 2014; Redo et al., 2012). Generalized linear models and generalized linear mixed models (GLMMs), Bayesian networks, and artificial neural networks were compared for the prediction of deforestation in forests in Mexico and Madagascar (Mayfield et al., 2017), and Spatio-temporal Bayesian Network approaches evaluated deforestation risk analysis in Brazil (Silva et al., 2020).
Lately used Machine Learning techniques such as Random Forest (RF) have shown very good accuracies, and in some cases, some advantages compared to previous methods (Breiman, 2001). RF has been implemented for the assessment of deforestation trends and its main drivers in Bolivia (Redo et al., 2012), for predicting future deforestation risk in Borneo (Cushman et al., 2017), and for the spatial prediction of deforestation probability in India (Saha et al., 2020). RF has also shown to be accurate in other ecological disciplines involving forests, such as prediction of tree species presence in the United States (Evans and Cushman, 2009), prediction of forest loss due to wind damage in Southern France (Hart

et al., 2019) or susceptibility to landslides in protected and non-protected forests in Iran (Shirvani, 2020).

RF, when used in predictive models, is capable of identifying complex interactions between variables, especially when the response-predictor relationships are non-linear and change spatially (Stan and Sanchez-Azofeifa, 2019; Zanella et al., 2017). It also provides better spatial accuracy compared with other models (Prasad et al., 2006) and has proven to be less sensitive to the removal of variables in comparison with other algorithms (Hart et al., 2019). It is a powerful tool with the ability to estimate the relative importance of the predictive variables, and also reduces the risk of overfitting, which is that the model is good at predicting the data used for training, but performs worse with independent test data (Willcock et al., 2018). Additional features of the use of RF are that it can provide spatially explicit prediction probabilities, which can be extremely useful for decision-making (Saha et al., 2020).

Most of the studies that used RF as a predictive tool for deforestation risk found that biophysical and climate factors were more important than socioeconomic ones and that accessibility, distance to markets, and topography were always among the most influential factors driving deforestation. Although this statement can be questioned, since accessibility and distance to markets are human-induced factors, but they are controlled by biophysical ones (Aide et al., 2013; Cushman et al., 2017; Redo et al., 2012; Saha et al., 2020; Zanella et al., 2017).

1.4. Monitoring forest cover and environmental services throughout time using remote sensing: The topography factor

Remote sensing (RS) and Geographic Information Systems (GIS) techniques are essential tools that have increasingly been employed in Costa Rica and worldwide to monitor forest cover changes and the associated changes in environmental services provided by forests (De Araujo Barbosa et al., 2015; Sader and Joyce, 1988; Stan and Sanchez-Azofeifa, 2019; Vallet et al., 2016). Current trends in satellite imagery, its free availability, and wide access to large-scale cloud computing like the Google Earth Engine platform assure that the use of these techniques will be increasing (De Araujo Barbosa et al., 2015; Gorelick et al., 2017).

Vegetation indices (VI), defined as the arithmetic combination of two or more bands related to the spectral characteristics of vegetation (Liu and Huete, 1995; Rouse et al., 1973), have been used in a variety of fields including phenology, classification of vegetation, photosynthetic activity, aboveground net primary productivity and land surface temperature (Cao et al., 2016; Liu et al., 2020). Vegetation indices, particularly the Normalized Difference Vegetation Index (NDVI), are essential components of any study aiming to investigate environmental services especially those where vegetation, water, and biodiversity are involved (Cord et al., 2017; De Araujo Barbosa et al., 2015).

However, VI sensitivity is affected by the changing radiance that accompanies changes in orientation of the vegetation surface being sensed (Matsushita et al., 2007). The radiance changes at different times in the year and between years, due to different solar incidences over the surface, the so-called sun-sensor geometry (Teillet et al., 1982). Radiance is further changed in rough terrain, where a combination of the orientation of the terrain and the position of the satellite will determine high or low illumination conditions (IC).

The Enhanced Vegetation Index (EVI), in turn, is more sensitive than NDVI to biophysical attributes such as the Leaf Area Index (LAI) (Galvão et al., 2016; Peng et al., 2018) and much more affected by IC than NDVI, because is not a ratio-based VI and cannot compensate for variations in IC. In addition, EVI was proven to be five times more sensitive than NDVI to changes in Near-Infrared reflectance (NIR) (Galvão et al., 2016; Maeda et al., 2014; Maeda and Galvão, 2015; Peng et al., 2018).

Due to the effect of IC, the use of EVI as an indicator of vegetation functioning or forest productivity has been put under debate in recent years. Some authors claimed that the unusual greening effect observed in the dry season in the Amazon forest using EVI Moderate-resolution Imaging Spectroradiometer (MODIS) imagery (Huete et al., 2006) was induced by changes in sun-sensor geometry and not as a result of canopy structure, phenological patterns, or vegetation functioning (Morton et al., 2014). This effect has been confirmed in similar ecosystems at different times of the same season, with different IC (Galvão et al., 2011; Maeda and Galvão, 2015). However, in similar tropical forests, after removal of the IC effects, seasonal patterns were still present and seemed to be correlated with gross primary production (GPP), although authors recommended being cautious with this correlation (Maeda et al., 2014). Effects related to sun-sensor geometry and topography

were also found at different times of the year using EVI and NDVI in subtropical deciduous forests under different IC. Sunlit and shadowed surfaces showed respectively different intensities of decrease and increase in reflectance even after topographic correction (Galvão et al., 2016).

In Costa Rica, the most valuable cloud forests in terms of biodiversity and PWAs are frequently located in low accessible areas, at high altitudes and irregular terrain (Bernard et al., 2009; Pfaff et al., 2009b). It is expected that topography will affect VI and consequently the evaluation of forest cover change or the environmental services derived from the calculation of these indices.

2. Research objectives and main research questions

This research aims to examine deforestation patterns in the cloud forests of Costa Rica and understand the biophysical and socioeconomic drivers behind them. Based on this analysis, develop a predictive model of deforestation risk that can be used to improve additionality and efficiency in the design of forest conservation schemes or policies such as the program of payments for environmental services. Additionally, it will examine the use of remote sensing techniques throughout time to evaluate their use as a tool to monitor forest cover and its environmental services associated.
In this sense, the four following specific objectives were established, namely:

1) To examine historical deforestation trends in the area and describe main biophysical and socioeconomic drivers
2) To use historical deforestation trends to develop a spatially explicit model of deforestation risk
3) To evaluate the effect of existing forest protection policies on historical and predicted deforestation trends and discuss their efficiency and additionality
4) To examine the use of remote sensing to monitor forest cover and the provided environmental services in the long-term

The main research questions to respond to these objectives is:

- How biophysical and socioeconomic factors influence historical and predicted deforestation?
- How the historical and predicted deforestation reveals the efficiency and additionality of PES?

3. Methods

3.1. Study area

The study area comprises the high part of the Reventazón basin in the province of Cartago, Costa Rica, with 176,192 hectares (Figures 1 and 2). It is approximately centered at 9.82ºN 83.67º W. This part includes the main cities of Cartago and Turrialba and other municipalities. The topography is mountainous and elevation ranges from 233 m to 3,461 m above sea level in Cerro de la Muerte, which is the highest road location in all Central America. Climate is tropical humid with average rainfall between 1,500 and 8,000 mm/year (Bernard et al., 2009). Principal land uses are agriculture, grassland for dairy cattle, coffee, sugar cane, and an important portion of secondary and primary forest (Figure 3). The basin is also important for producing 25% of the hydroelectric power of the country (Vallet et al., 2016). The lower part of the Reventazón basin, which extends downstream the Reventazón dam inaugurated in 2016 was not included in this study, it is well-differentiated in climate, land, and vegetation with a Caribbean influence and different land cover and management.

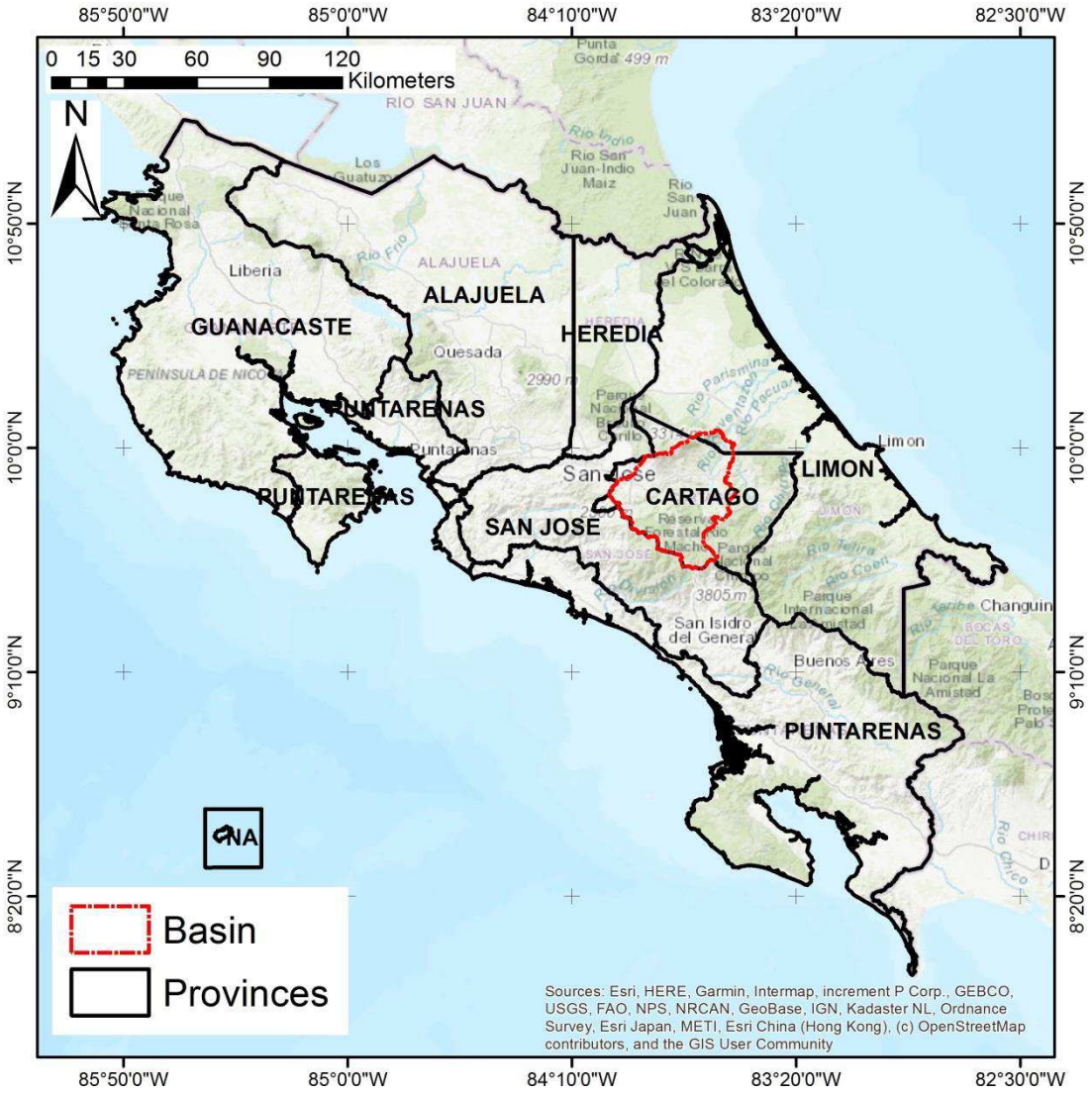


Figure 1. Location of the study area in Costa Rica. The high part of the Reventazón basin is located in the Central Valley, in the Cartago province.

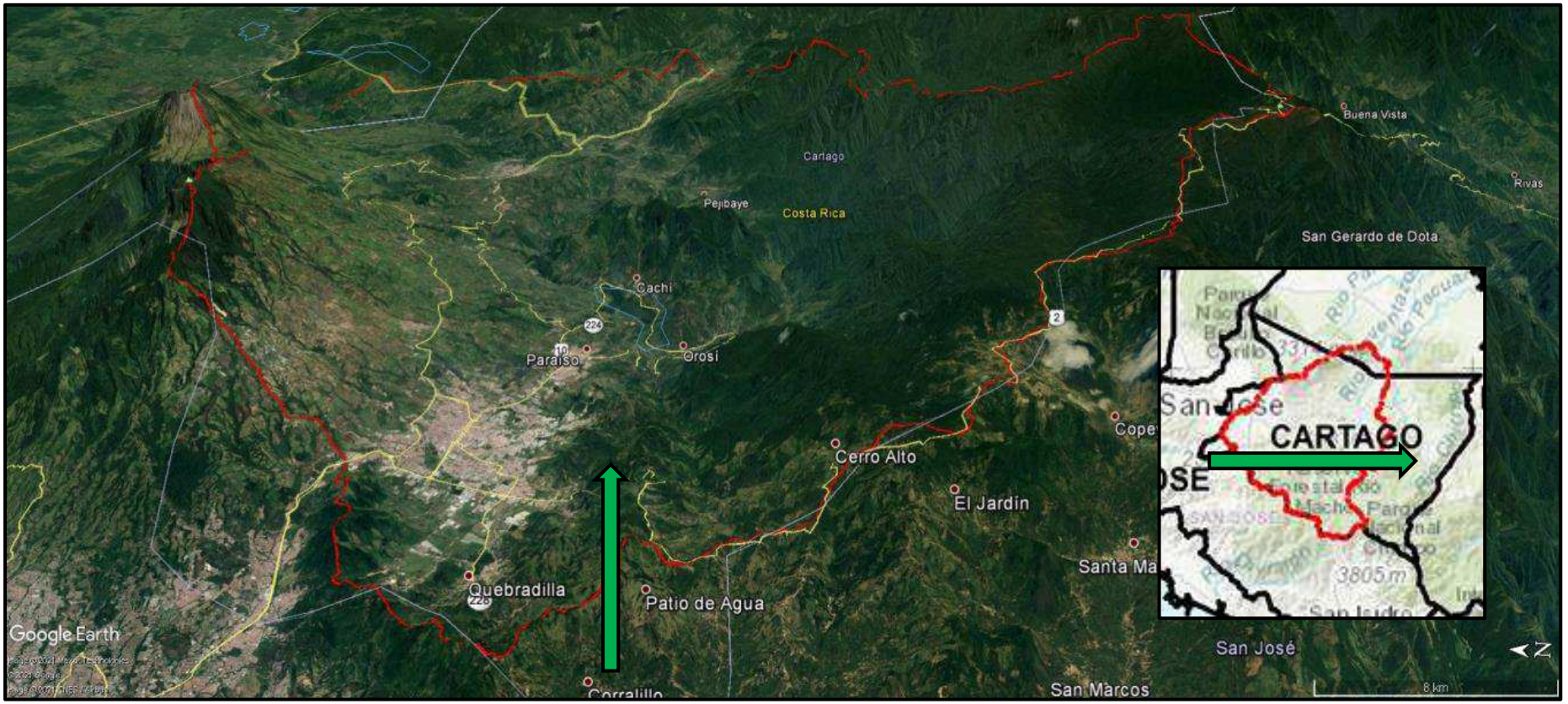
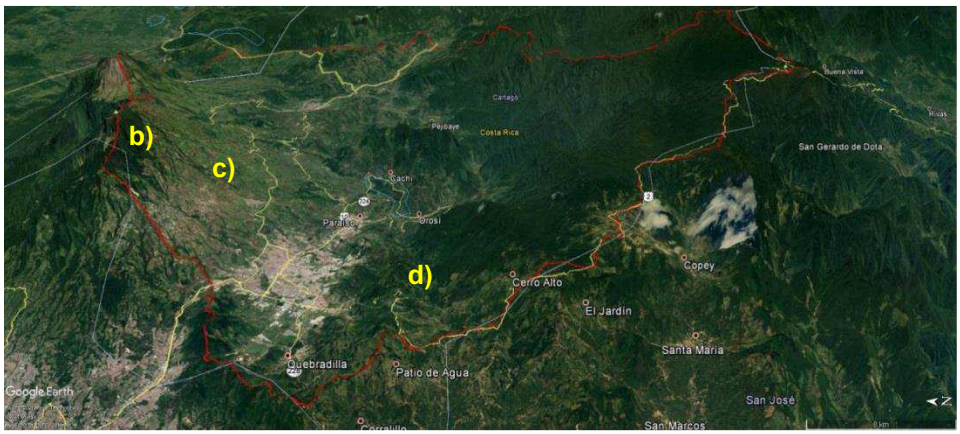
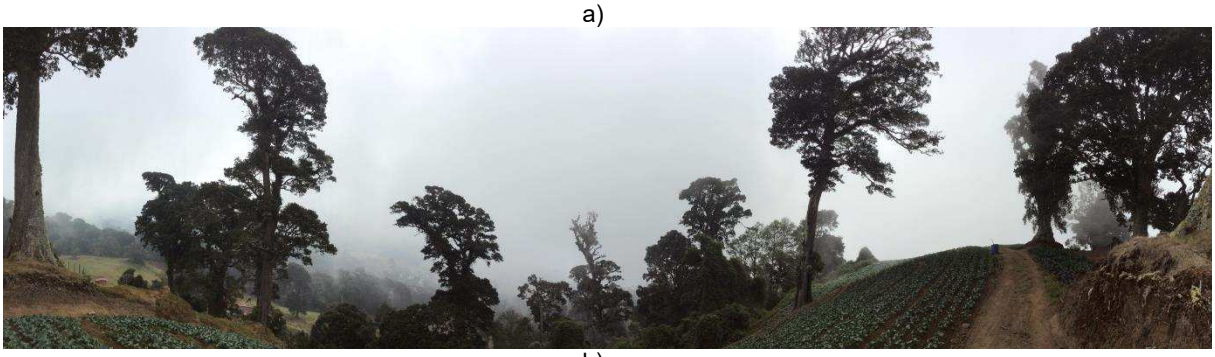


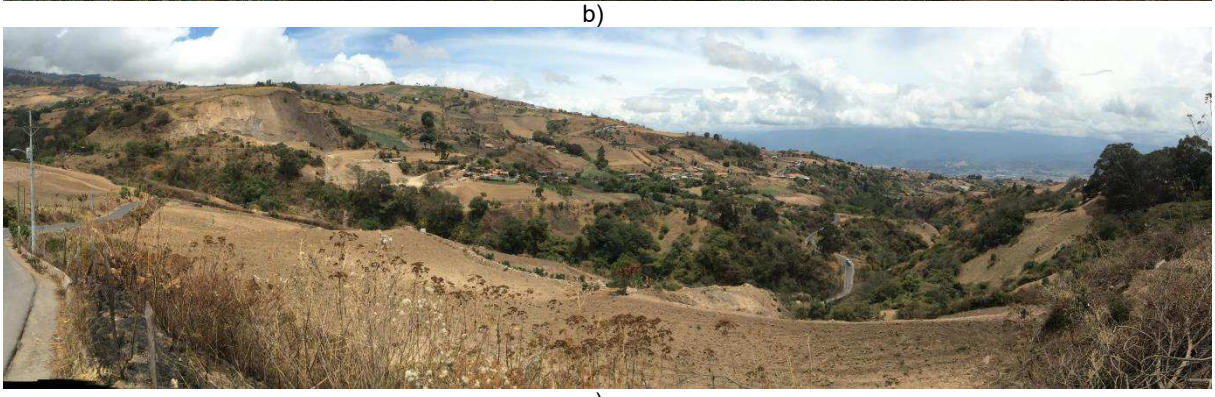
Figure 2. View of the high part of the Reventazón basin from West to East using a 3D model, the green arrow in the right miniature indicates the perspective of view. The terrain has been exaggerated using Google Earth. The shape of the Central Valley is seen. The city of Cartago in the foreground and main roads run through the valley floor, ending in the Caribbean coast. To the left side Volcanoes Irazú and Turrialba and the agricultural areas and pastures at their slopes. To the right side the mountainous terrain is the home of the cloud forests and the Tapantí national Park, at its top the highest point of the basin is found: Cerro de la Muerte (3,461 m). Source: Google Earth.



a)



b)



c)

d)

Figure 3. Panoramic pictures representative of the landscapes of the study area. a) Overview of the basin with the approximate location of the pictures. b) Scattered trees of mature cloud forest above agricultural areas close to the Irazú crater. c) Agricultural areas and pastures at the slopes of Irazú volcano. d) Valley of Orosi, in the background of the picture the cloud forest Tapantí National Park.

3.2. Biophysical approach

3.2.1. Protected wildlife areas

There are 17 protected wildlife areas (PWA) in the study area which account for 42.06% of the total basin area, including national parks, forest reserves, and other types of designated areas (Table 2 and Figure 4). National parks and biological reserves are the most restrictive, and no land-use change is allowed inside them (Pfaff et al., 2009a). Spatial limits for PWA were extracted from the Atlas of Costa Rica (2014) in vector format. The extent of some of them was reduced and clipped using the border of the basin as the frontier. The rest of the area outside the PWA was labeled as "Basin" and was also included in the analysis. Historical vegetation loss and predicted vegetation loss risk are evaluated inside and outside all the PWA in the basin as a measure of the effectiveness and additionality of establishing these protected areas.

Table 2. Characteristics of the Protected Wildlife Areas in the Study area

Category	Name	Area (ha)	Declaration year
Forest reserve	C.V.Central	4,299.93	1975
	Los Santos	168.30	1975
	Río Macho	4,170.06	1964
Protective zone	Carpintera	793.17	1976
	Rio Tuis	3,699.09	1986
	Navarro-Sombrero	6,200.28	1984
National Park	Chirripó	218.25	1975
	Los Quetzales	7.02	2006
	Tapantí	42,499.71	2000
	Volcán Irazú	1,051.29	1955
	Volcán Turrialba	405.9	1955
Biological reserve	Cerro Vueltas	8.82	1994
Private National Wildlife Refugee	Steve Friedman	19.35	--
	La Marta	1,267.38	--
Wetland	Bonilla-Bonitilla	10.26	1994
Natural Monument	Guayabo	213.84	1973
State property	F.estado	2.16	--

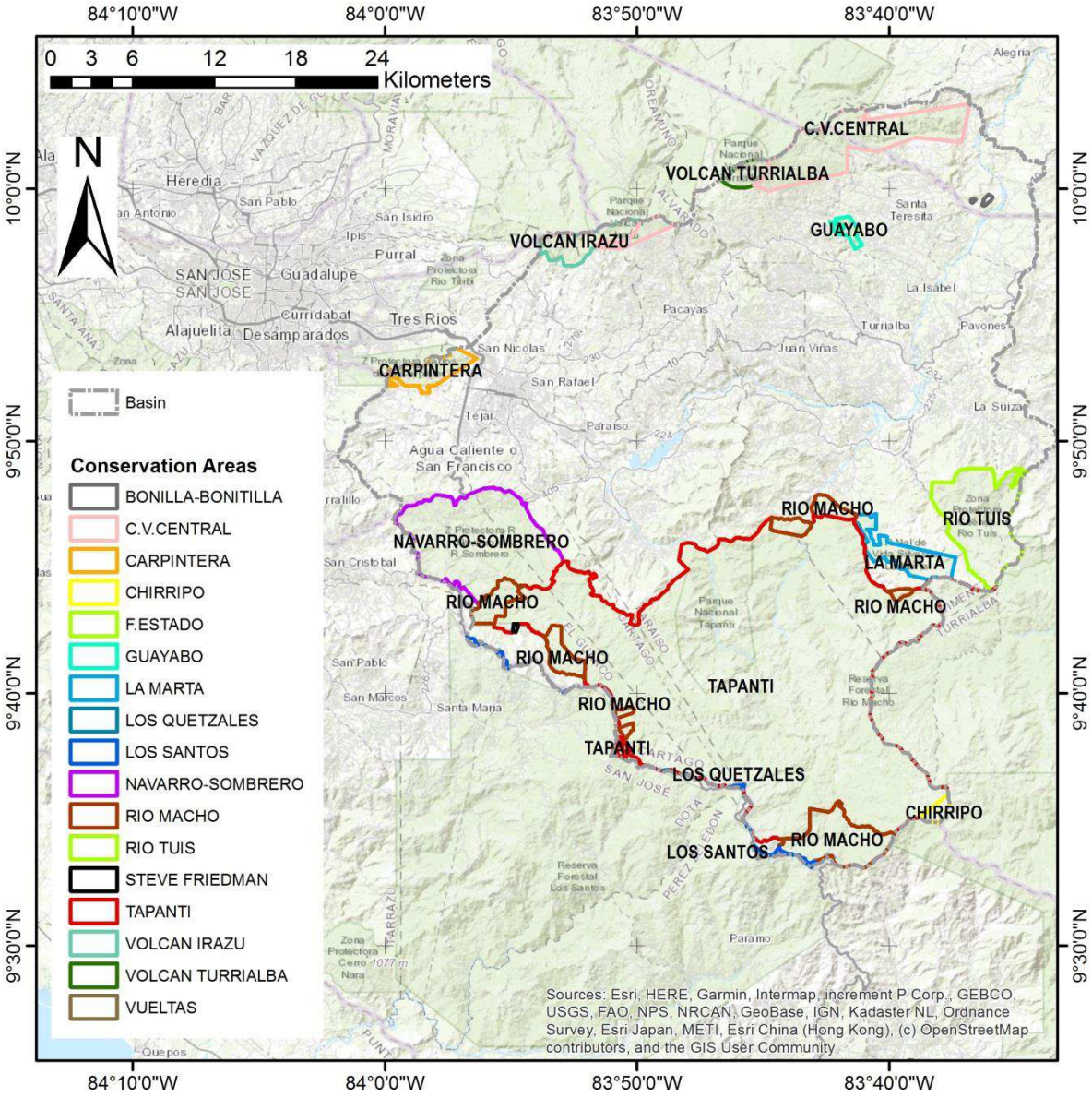


Figure 4. Location of the Protected Wildlife Areas inside the Reventazón basin

3.2.2. Private farms visited and delineated

Spatial data of 67 private farms were obtained through visits during the years 2016 and 2017 in the study area. Key informants from research institutions and local government offices of environment and agriculture provided the first contacts from farmers. After visiting an initial number of them, contacts of new farmers contacts were made. Some of the property maps were shared by the farmers and in some cases digitized maps were available in vector format such as shapefile or using National cadastral maps available

online. When visiting a property, a sketch of the different land uses was made and geographic coordinates of the location were collected. When network was available, Google Maps® in the smartphone was used to better locate the boundaries and reference objects. Back in the office, the farm was divided and delineated manually in the different land uses using Google Earth® and high-resolution imagery available, frequently Maxar® or DataGlobe® at less than 1-meter pixel resolution. ArCGIS® Desktop 10.8.1 was also used when needed in the digitization process.

Land uses in the area were grouped into 7 categories: Agriculture (potatoes, onions, carrots, and other vegetables), sugar cane, coffee, forest, infrastructure (roads, buildings, greenhouses, water reservoirs), pasture, and pasture with trees. Land uses were frequently clearly divided within the farm and in homogeneous and continuous portions (Figure 5). The forest category did not distinguish between planted or natural or secondary forests. The only condition to consider a continuous forest patch was that trees should be closer to 30 meters. This assumption is related to the spatial resolution of Landsat sensors which is 30 meters and has been intensely used in this research. It is expected that trees if very isolated will not be differentiated from the land cover underneath, i.e., pasture (Bastin et al., 2017).

3.2.2.1. Biophysical differences between PES and non-PES farms

It is assumed that biophysical variables between PES and non-PES farms may differ to some extent and that is possible therefore to find statistical differences considering both categories of farms as treatment groups. Biophysical variables are shown in section 3.2.3.2, from these variables, a selection of some of the most important variables for the predictive model was extracted for statistical analysis, namely: distance to roads, distance to productive land use, slope, aspect, elevation, and NDVI. The values of the biophysical variables in PES and non-PES farms were analyzed using descriptive statistics, such as mean, median, and standard deviation. Boxplots and violin plots were also employed for better visualization of data distributions. The distributions of all variables were evaluated using the Anderson-Darling test for normality (Shapiro-Wilk test has a limit in R when n> 3,000). Since the distributions of most of the biophysical variables were found to violate the assumption of normality and homogeneity of variance (Hollander et al., 2014), the Wilcoxon signed-rank test was applied, to check which variables were significantly different between PES and non-PES farms. The null hypothesis was that the median value

of the biophysical variables would be the same between PES and non-PES farms. All the statistical data analysis was carried out in R version 4.0.0 (https://cran.r-project.org/).

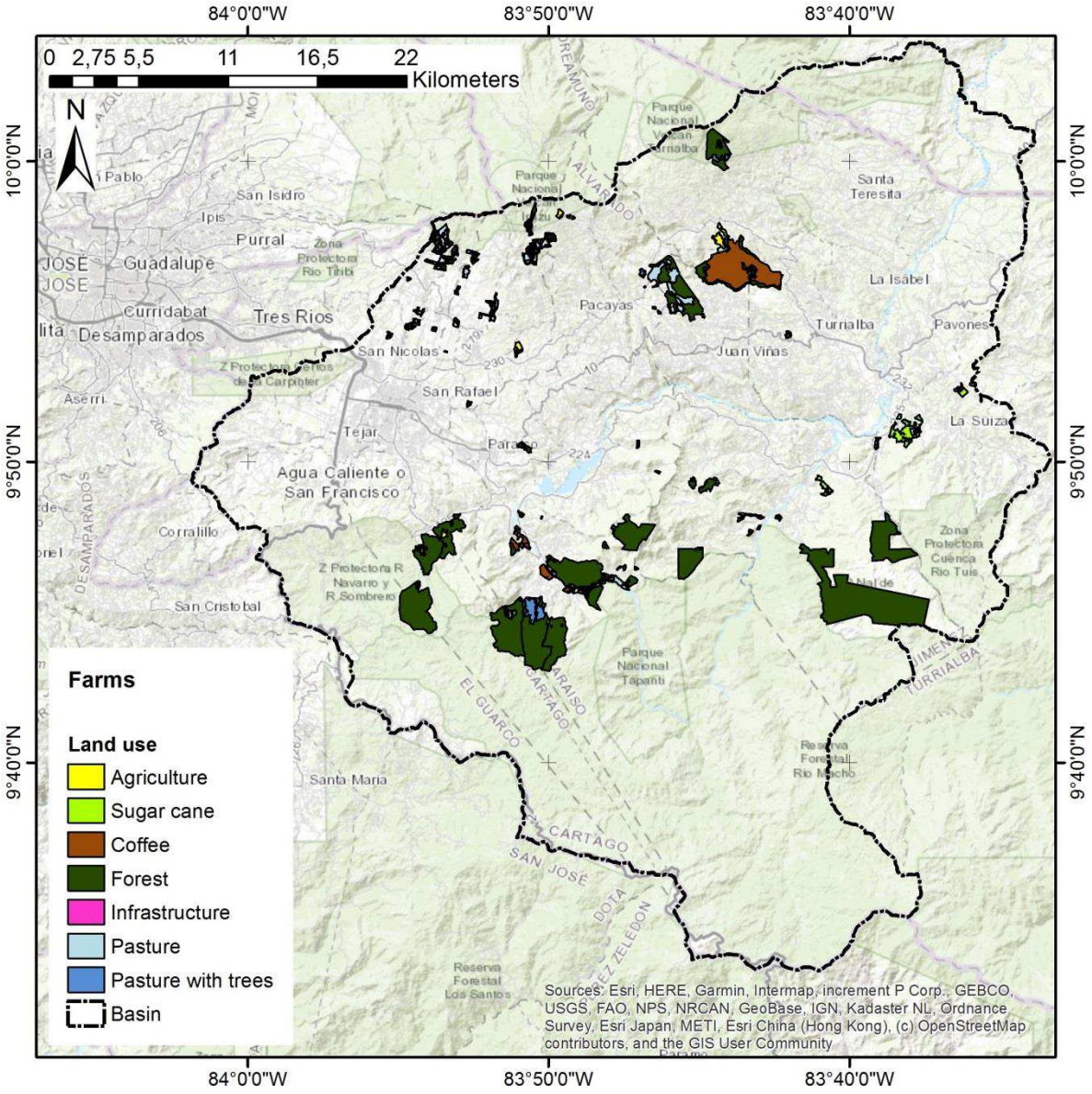


Figure 5. Location of the 67 farms in the study area and delineation of the different land use categories inside them.

3.2.3. Building a predictive model of vegetation loss risk

3.2.3.1. Selection of historical vegetation cover and vegetation loss data

The high part of the Reventazón basin has a total area of 176,192 hectares which includes all types of land cover and land use. From the total area only information on vegetation cover and vegetation loss was extracted from the dataset: ee.Image("UMD/hansen/global_forest_change_2018_v1_6") available at Google Earth Engine (GEE). This dataset provides spatial information of tree cover at 30-meter pixel resolution using Landsat sensors and also spatial information where forest loss has occurred throughout the years 2000 to 2018 (Hansen et al., 2013).

Datasets can also be downloaded at: https://storage.googleapis.com/earthenginepartners-hansen/GFC-2020-v1.8/download.html.

In the year 2000, according to this dataset, 88% of the Reventazón basin (154,983 hectares) had a vegetation cover higher or equal than 1% in tree cover on a 30 by 30-meter square (pixel) and taller than 5 meters in height (Hansen et al., 2013). A minimum threshold for vegetation cover or forest definition was not selected as 90% of the data has a tree cover above 44%, so it is considered that most of the area was forested in the year 2000 (Figure 6).

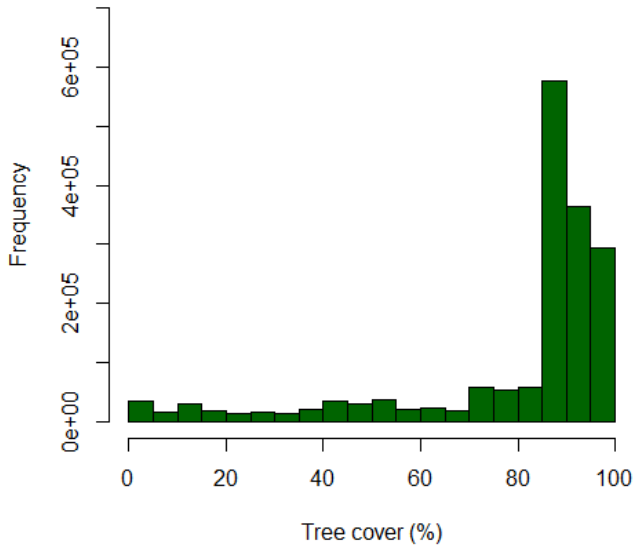


Figure 6. Distribution of tree cover for the selected vegetation cover image for the year 2000.

Vegetation loss data throughout the years 2000 to 2018 in this dataset was selected to spatially identify the characteristics of the areas where vegetation loss occurred in the last 18 years. Vegetation loss data has been used in other studies as a source of historical deforestation trends (Bax and Francesconi, 2018; Dlamini, 2016; Goldman et al., 2017).
A portion of 2.52% of the study area vegetated in the year 2000 (4,439.88 hectares) has been lost between the years 2000-2018. Following these numbers an average of 246.66 hectares was lost per year which accounts for a year loss rate of 0.14%, indicating a smooth decreasing trend in vegetation loss (figures 7 and figure 8).

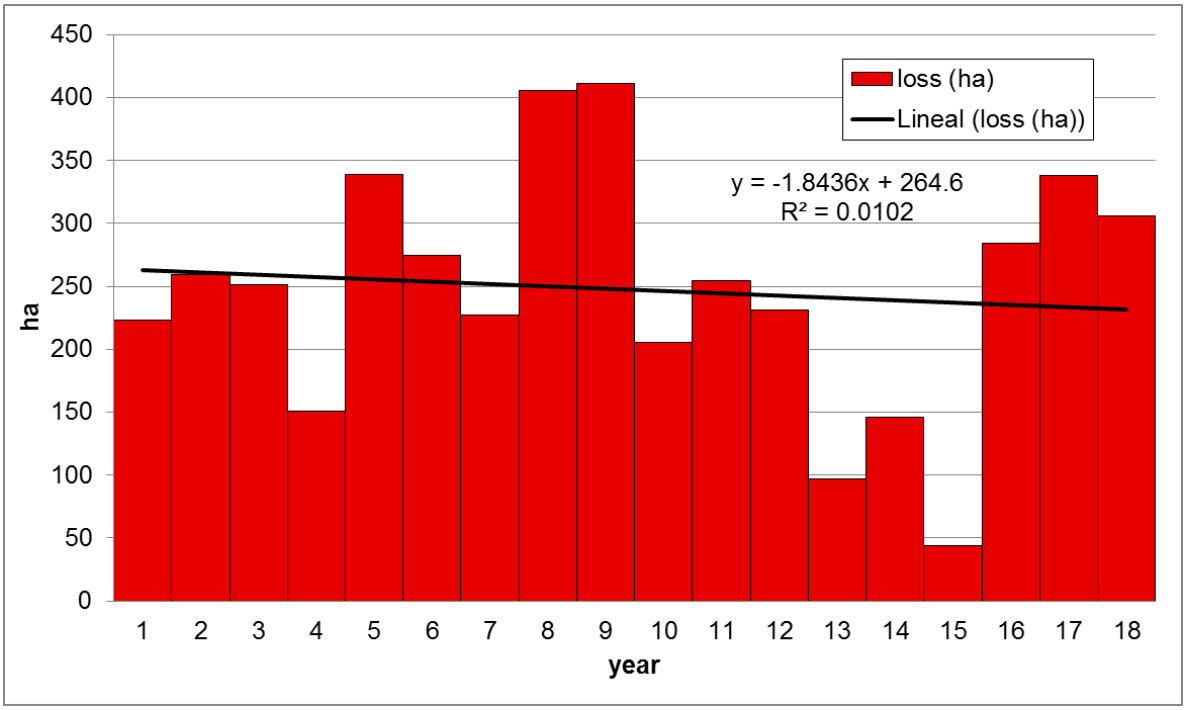


Figure 7. Vegetation area lost in hectares in the study area by year since the year 2000 and the linear trend.

Pixels with a vegetation cover <1% were discarded from the analysis because based on visual interpretation these comprise mainly urban areas infrastructures, water, and bare land. A total of 49,332 vegetation loss pixels were detected during the period 2000-2018. A similar number of no-loss pixels (n=49,232 pixels) were calculated in the remaining part of the Reventazón basin (Bax and Francesconi, 2018) using a random spatial approach and ensuring a minimum distance of 30 m to avoid spatial coincidence between loss and no loss pixels**.** To employ this information in the machine learning model, loss pixels were labeled as 1 and no loss pixels as 0.

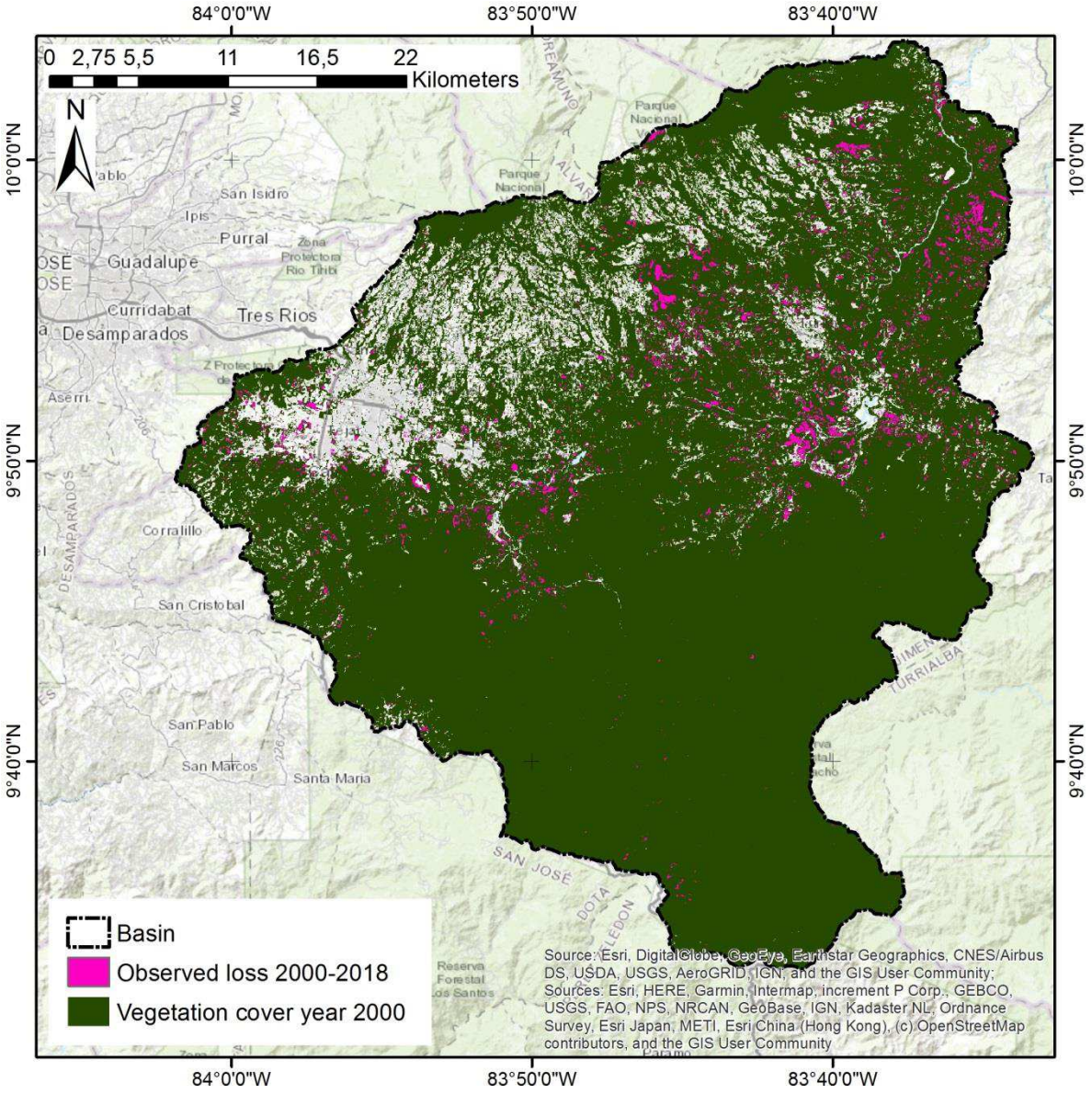


Figure 8. Observed historical vegetation loss over the study area during the years 2000-2018.

3.2.3.2. Predictor variables used to train the model

Several physical features that have been shown to affect the risk of tree loss or deforestation were selected (Aguilar-Amuchastegui et al., 2014; Cushman et al., 2017; Evans and Cushman, 2009; Saha et al., 2020; Silva et al., 2020), Table 3 and Figures 9 to 13. These included physical properties of the land cover itself, which can be revealed using multispectral remote sensing data such as Landsat. The Landsat mosaic with surface reflectance values for the year 2000 was used for this purpose and it was available in the same Hansen product in GEE from where the vegetation loss data was extracted.

Surface reflectance has proven also to be a good tool to detect subtle spectral changes that indicated logged areas (Hethcoat et al., 2019). Several known indices were calculated using combinations of different available bands in the Landsat mosaic. These indices have demonstrated the capability of separation between different land cover classes and vegetation, between croplands and forests, and forest and plantations (De Alban et al., 2018).

Table 3. Spatial datasets used as predictor variables in the model

Name of band/index/measure	Description	Units
RED	Landsat 7 band 3 (RED) cloud-free image composite. Reference multispectral imagery from the first available year, typically 2000.	0.63-0.69µm
NIR	Landsat 7 band 4 (NIR) cloud-free image composite. Reference multispectral imagery from the first available year, typically 2000.	0.77-0.90µm
SWIR1	Landsat 7 band 5 (SWIR) cloud-free image composite. Reference multispectral imagery from the first available year, typically 2000.	1.55-1.75µm
SWIR2	Landsat 7 band 7 (SWIR) cloud-free image composite. Reference multispectral imagery from the first available year, typically 2000.	2.09-2.35µm
Normalized Difference Vegetation Index (NDVI)	(NIR - RED)/(NIR + RED) Widely used vegetation index to identify vegetation cover (Rouse et al., 1973)	-1 to 1
Land Surface Water Index (LSWI)	(NIR - SWIR1)/(NIR + SWIR1) The vegetation index is sensitive to liquid water in the vegetation (Gao, 1996)	-1 to 1
Normalized Difference Tillage Index (NDTI)	(SWIR1 - SWIR2)/(SWIR1 + SWIR2) The vegetation index is sensitive to different agricultural management practices (tillage, drainage) and soil properties as soil texture (Van Deventer et al., 1997)	-1 to 1
Soil-Adjusted Vegetation Index (SAVI)	SAVI = [(NIR - RED)/(NIR + RED + L)] * (1+L). Vegetation index that minimizes soil brightness influences below dense canopies (Huete, 1988)	-1 to 1
Modified Soil-Adjusted Vegetation Index ($MSAVI_2$)	$[2 * NIR + 1 – [(2 * NIR + 1)^2 - 8 * (NIR – RED)]^{1/2}]/2$ Modified SAVI that automatically adjusts its L values to optimal and has the ability to further decrease soil brightness (Qi et al., 1994)	-1 to 1
Soil-Adjusted Total Vegetation Index (SATVI)	((SWIR1 - RED)/(SWIR1 + RED * 0.1)) * (1.1 - (SWIR2/2)) Vegetation index sensitive to herbaceous vegetation cover, height, and biomass (Marsett et al., 2006)	-1 to 1
Elevation	Shuttle Radar Topography Mission (SRTM) is a digital elevation model at 30-meter pixel resolution (Farr et al., 2007)	Meters
Slope	Slope is derived from the previous SRTM at 30-meter pixel resolution	Degrees
Aspect	Aspect is derived from the previous SRTM at 30-meter pixel resolution	Degrees
Distance from main roads	Derived 1-meter pixel resolution from the road shapefile available at Atlas de Costa Rica (2010) (https://www.tec.ac.cr/)	Meters

The rest of the variables included were related to topography factors such as elevation, slope, or aspect. A raster image showing distance to the nearest road was also produced using the vector layer available in the “Atlas of Costa Rica 2010”. Distance to roads is an important factor to predict deforestation risk (Cushman et al., 2017; Silva et al., 2020; Valle et al., 2020).

Values for a total of 14 predictor variables were extracted for each class (0=no loss, 1=loss), all this information was used to train the Random Forest classifier. All variables were resampled at a pixel resolution of 30 m.

In this study, it is assumed that future vegetation loss would follow past patterns of change and that the prevailing conditions and factors that facilitated deforestation would not change significantly (Aguilar-Amuchastegui et al., 2014).

Finally, socioeconomic predictors were excluded, although these may be the hidden cause behind some of the physical changes observed in the vegetation. Some authors have found that its inclusion can have little effect in models for predicting deforestation (Geist and Lambin, 2002; Redo et al., 2012).

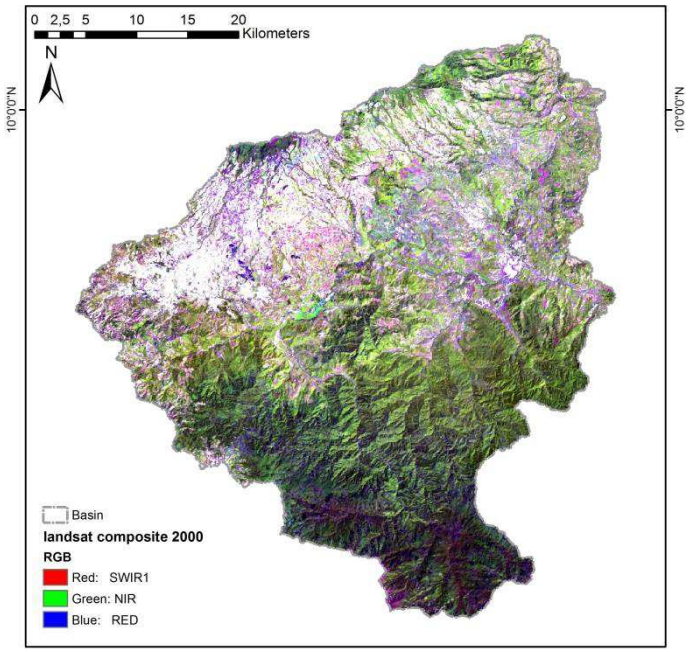


Figure 9. Landsat composite of the study area for the year 2000. The combination of bands are SWIR1, NIR, RED. Dark green areas indicate mature forest, clear green to bright green secondary/young forests to healthy crops. Pink and purple areas correspond to pastures/crops. White areas are not included areas because the vegetation was <1% cover (infrastructures, buildings, and water)

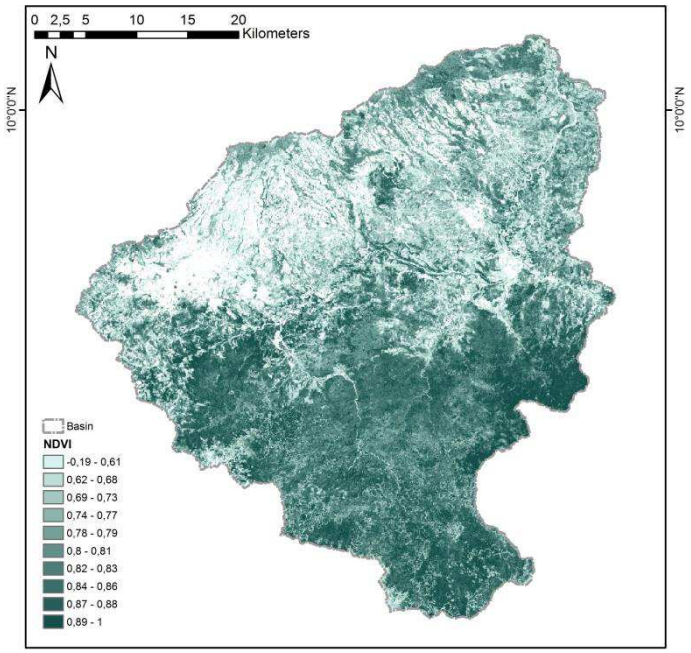


Figure 10. NDVI image of the study area for the year 2000. Increasing vegetation cover from light to dark green.

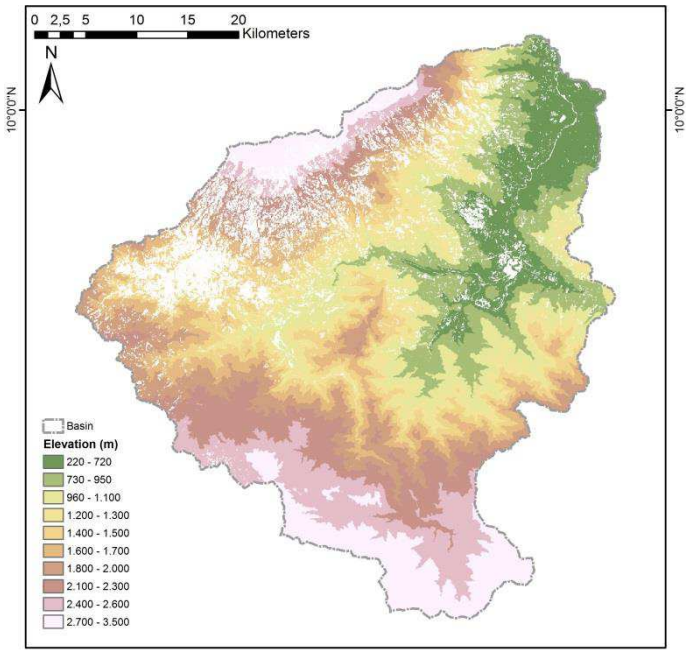


Figure 11. The elevation map of the study area for the year 2000.

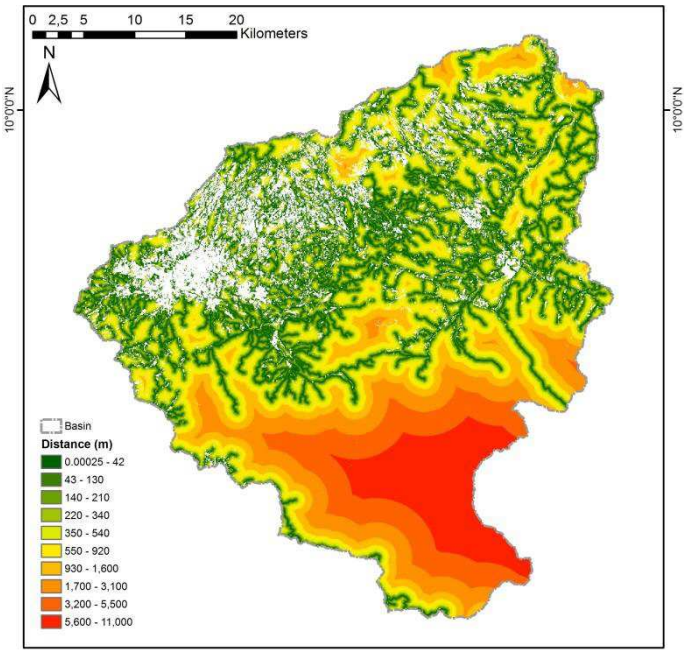


Figure 12. Distance from main roads map of the study area for the year 2010.

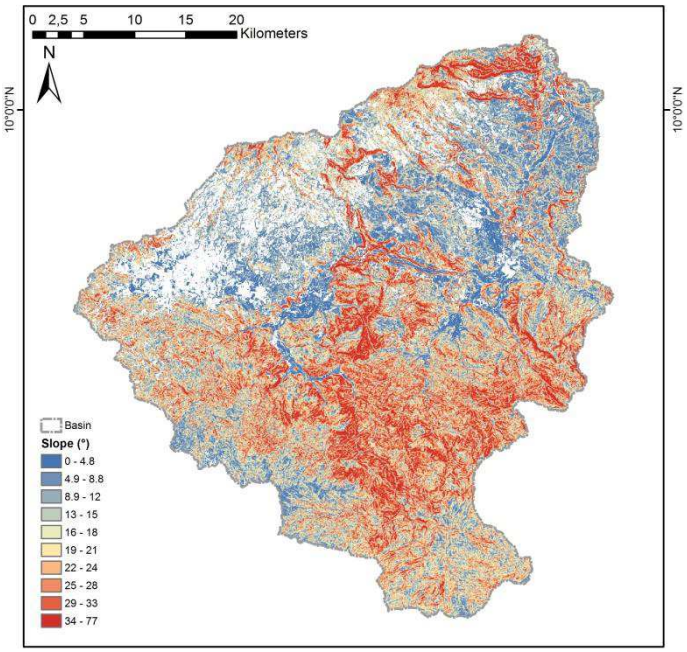


Figure 13. Slope map of the study area for the year 2000.

3.2.3.3. Random Forests to model vegetation risk loss

The risk of potential vegetation loss was predicted using the Random Forest method (RF) (Breiman, 2001). It has recently become a standard non-parametric classification tool for constructing prediction rules based on predictor variables without making any prior assumption or association with the response variable (Probst et al., 2019). RF is a machine learning approach that creates multiple decision trees to execute a classification (Saha et al., 2020) with enough accuracy to be used as a prediction tool (Probst et al., 2019). The mode of these trees is issued to create an "ensemble" tree that makes a prediction (Mascaro et al., 2014). RF is also useful to identify complex interactive and non-linear relationships between the predictor and response variables (Zanella et al., 2017).

All available pixels with the associated information on predictor variables for classes 0 and 1 (0=no loss, 1=loss), were merged into a single multipoint vector layer (n=98,564). A random number between 0 and 1 was given to each of these points to split the data into training and testing points. 70% of these data were used to train the RF model and 30% were not included and were left aside for testing. To ensure greater spatial variability and avoid spatial autocorrelation between training and testing pixels, the data were again given a random number between 0 and 1, and 70% of each training and testing group were finally selected (De Alban et al., 2018).

Random Forest has its internal procedure to calculate the error in the prediction. It randomly selects a portion of the data for model training and reserves a portion for model testing, the so-called out-of-bag (OOB). This portion of OOB predictor variables is used to test the proportion of misclassification (%), which is a measure of the error of the model in predicting the output variable and is called OOB error rate (%) (Humphries et al., 2018).

To build the random forest, the hyperparameters number of trees (trees), randomly drawn candidate variables (mtry) and node size (n) are required, and different combinations of these are responsible to increase or decrease the OOB error rate. The "tuneRanger" package in R was employed to find the optimal hyperparameters which help to achieve the higher performance in terms of finding the lowest OOB error rate (Probst et al., 2019).

Once the optimal hyperparameters were found, the "ee.Classifier.randomForest" algorithm in GEE was used to generate the classified raster image of vegetation loss risk.

3.2.3.4. Validation method

Besides the internal validation test that RF provides using the OOB error rate, the predictions were validated using testing data that were not used in the training of the RF model. The receiver operating characteristics (ROC) curve is a well-accepted method of validation (Fawcett, 2006). It is a graphical presentation of the true positive rate (TPR) or sensitivity, corresponding to correctly predicted absences or proportion of true positives against the false positive rate (FPR) or specificity corresponding to correctly predicted presences or proportion of false positives (Drew et al., 2010; Humphries et al., 2018).

The balance between sensitivity and specificity is an indication of how well the model performs. The Area under the ROC Curve (AUC) value ranges from 0 to 1, where 0.5 indicates a weak prediction precision of the model or no significant difference from randomness and 1 indicates perfect classification or prediction (Drew et al., 2010). Values above 0.8 are considered very good in ecology (Humphries et al., 2018).

3.2.3.5. Variable Importance

RF for classification is also effective in ranking the importance of the predictor variables in explaining the predicted variable (Probst et al., 2019; Redo et al., 2012). Measures of variable importance are called Mean Decrease in Accuracy (MDA) and Mean Decrease in Impurity or Gini Impurity (MDI). MDA is defined as the contribution of a predictor variable to the accuracy of the model and measures how the permutation of this variable can decrease the accuracy of the model (Probst et al., 2019). MDI measures the Gini impurity which tells us what is the probability of misclassifying an observation. The Gini impurity measure is used when RF builds the trees to decide the optimal split and subsequent splits based on the available predictor variables that reduce impurity at its minimum. Then the lower the decrease in Gini, the lower the likelihood of misclassification. However, MDA will be considered more reliable in this study since MDI can be strongly biased when different types of variables with different intervals and values are used in the model (Strobl et al., 2007). Also, MDA is a more appropriate criterion for prioritizing predictor variables (Arabameri et al., 2019)

Variable importance plots are useful to graphically interpret the contribution of all variables to the predictive accuracy of the model. Additionally, partial dependency plots are useful to

explain the effect of the most important individual predictor variables and how their effect change as their original units change.

3.2.4. Analysis of the topography and illumination condition (IC)

A private property located in the study area was selected (9°46'31.17"N, 83°45'27.31"W, WGS 84). The vegetation is classified as an evergreen premontane rain forest bordering with Tapanti-Cerro de la Muerte, an important national conservation area (Figure 14). The altitude ranges between 986-1302 m.a.s.l. The property is devoted to conservation and ecotourism with 88% of the area covered by little disturbed native forest. The property was visited in August 2017. After screening the area for forest changes using aerial photo single frames available for free at https://earthexplorer.usgs.gov/ and high-resolution imagery at different years available at Google Earth Pro (Figure 15) the area was considered as an invariant forest spot to carry out this long term analysis.

Figure 14. The geographical location of the area of interest (AOI), the red polygon corresponds to the Landsat tile for path and row 15/53.

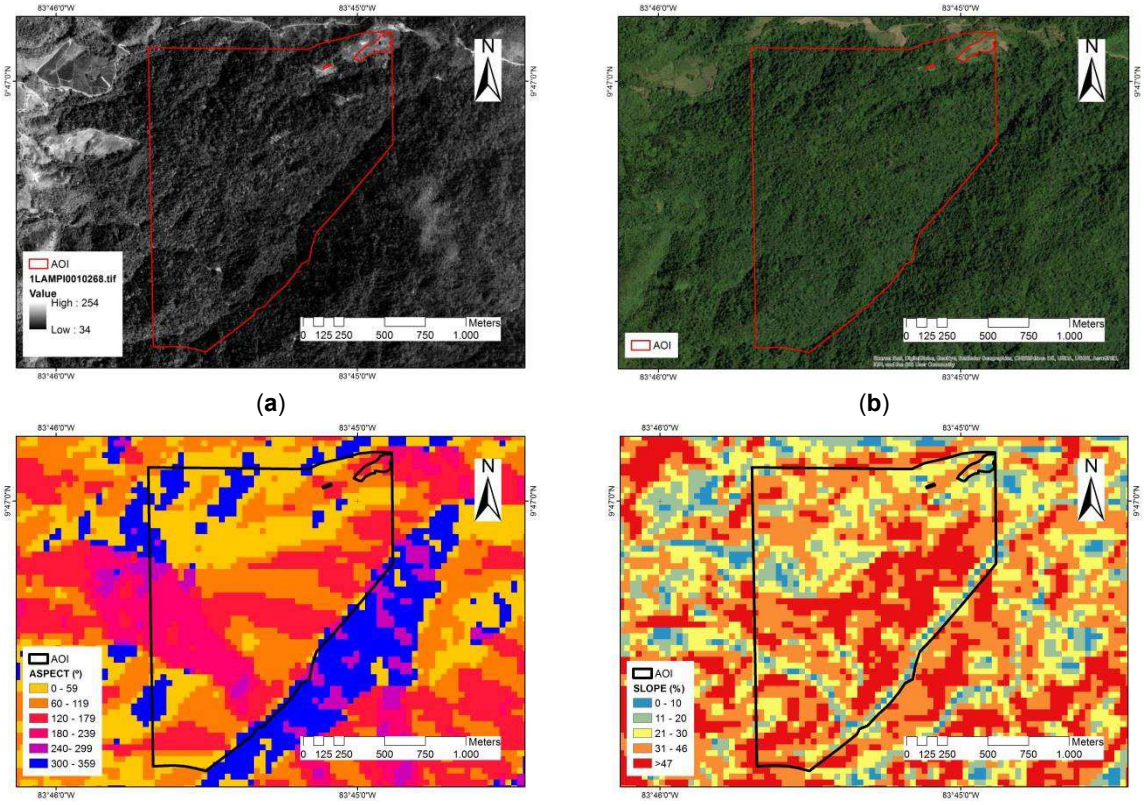


Figure 15. Area of Interest (AOI): (a) Aerial single frame of the area 11/08/1992, (b) Very high-resolution satellite image (16/01/2017) (DigitalGlobe), (c) Aspect (º), (d) Slope (%).

3.2.4.1. Landsat datasets and Image processing

Data were available and processed with the Google Earth Engine (GEE) platform. Landsat Surface Reflectance Tier 1 collections for Landsat 4 ETM, Landsat 5 ETM, Landsat 7 ETM+ and Landsat 8 OLI/TIRS were used in the analysis. Calibration between sensors has been done, so they are suitable for time-series analysis. Correspondent ImageCollection IDs in GEE are: 'LANDSAT/LT04/C01/T1_SR'; 'LANDSAT/LT05/C01/T1_SR'; 'LANDSAT/LE07/C01/T1_SR' and 'LANDSAT/LC08/C01/T1_SR'. The collections correct illumination/viewing geometry and atmospheric effects using LEDAPS in ETM sensors and LaSRC in OLI/TIRS sensors. (USGS, 2019a, 2019b). The collection was filtered for all images available during the period 01-01-1984 to 31-12-2017. Path and row for all images were 15/53.

For the removal of clouds, the CFMASK algorithm was used (Foga et al., 2017). After the removal of cloudy pixels in the area, clear and masked cloudy pixels are obtained for every image. GEE was used to count the highest number of contiguous clear pixels inside

the AOI, for the highest number of available images. This reduced the initial area to a limited number of pixels (n=1,228) (Figure 16). Finally, these pixels were cloud-free for 39 images (Figure 17). Every individual image was visually inpected to search for cloud or haze remnants.

Figure 16. Landsat pixels available (30x30m), n=1228 for the 39 selected images in the AOI after cloud masking, and screening.

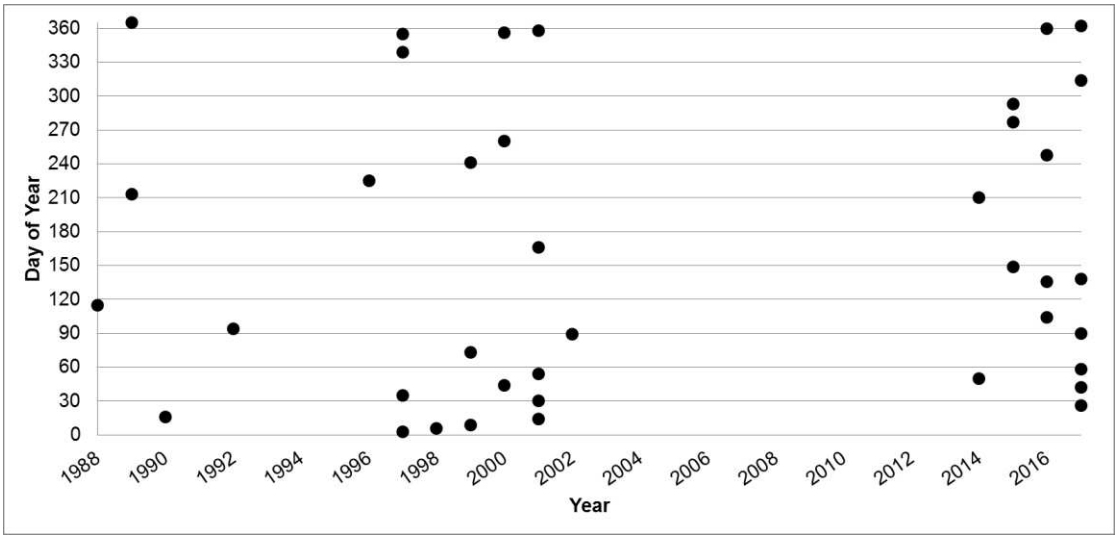


Figure 17. Dates expressed as year and day of year when the 39 selected images were available for the analysis.

Two common Vegetation indices (VI) were calculated for every image NDVI (Rouse et al., 1973) (Equation 1) and EVI (Liu and Huete, 1995) (Equation 2) and added as new bands.

$$NDVI = (NIR-RED)/(NIR+RED) \quad (1)$$

$$EVI = G*((NIR - RED)/(NIR + C_1*RED – C_2*BLUE + L)) \quad (2)$$

Where NIR, RED, and BLUE are Near-infrared; red and blue bands for Landsat images. G (2.5) is a scaling factor; L (1.0) is the canopy background adjustment factor; and C_1 (6.0) and C_2 (7.5) are the coefficients of the aerosol resistance term (Didan et al., 2015).

3.2.4.2. Illumination Condition (IC)

The incidence angle (i), is the angle between the normal to the pixel surface and the solar Zenith angle (z) (Holben and Justice, 1980; Smith et al., 1980) (Figure 18). The radiance detected by the sensor after interacting with surfaces of different slopes and aspects is dependent on this incidence angle (i). If (i) is low, the directions of the sun and the sensor are closer and radiance will be higher, the scene will be well illuminated and shadows due to topography minimized. Conversely, if (i) is large, the light will reach the surface in an oblique direction, illumination will decrease in the scene and more shadows will be cast due to the topography effect. Illumination condition (IC) is the cosine of the incidence angle (i) and ranges from 0 to 1, respectively, with 0 being values for poorly illuminated and 1 for well-illuminated areas. IC is calculated in Equation 3.

$$\cos i = \text{Illumination Condition (IC)} = \cos z \cos s + \sin z \sin s \cos (a – o) \quad (3)$$

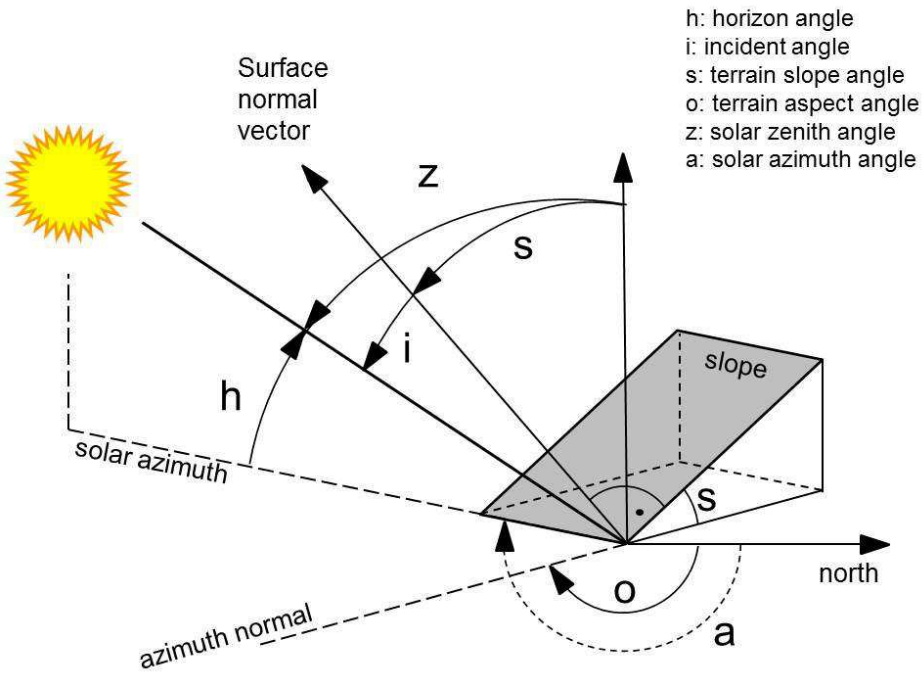


Figure 18. Terrain and solar angles involved in the calculation of the incidence angle (i) and illumination condition (IC).

Where z is the solar zenith angle; s is the terrain slope; a is the solar azimuth angle and o is the terrain aspect (Teillet et al., 1982). To calculate slope and aspect, a Digital Elevation Model (DEM), the Shuttle Radar Topography Mission (SRTM) was used, which is also available as a dataset in GEE (ID: 'USGS/SRTMGL1_003') (Farr et al., 2007). Zenith and azimuth angles are available in the metadata of each Landsat image. A script was created in GEE using SRTM and metadata solar angles in the image to calculate IC at the pixel level and incorporate it in each image as a new band.

3.2.4.3. Statistics across the collection of images

The 39 selected images, each one with the three bands IC, EVI, and NDVI (Figure 19b) were organized as a GEE Image Collection object. "Reducers" were used, which are a powerful tool in GEE that allow us to make computations at the pixel level across a stack of images or limited to a region in space (Figure 19a). The ee.Reduce algorithm was used to calculate descriptive statistics such as mean and standard deviation and correlations such as Pearson correlation. Mean IC and standard deviation IC were calculated at the pixel level for the 39 images. Mean IC shows which pixels are more frequently shadowed

or illuminated with respect to the terrain and the sensor. The standard deviation image indicates the magnitude of variation of the IC for each pixel. Finally, two images were obtained, a mean IC image, and a standard deviation IC image.

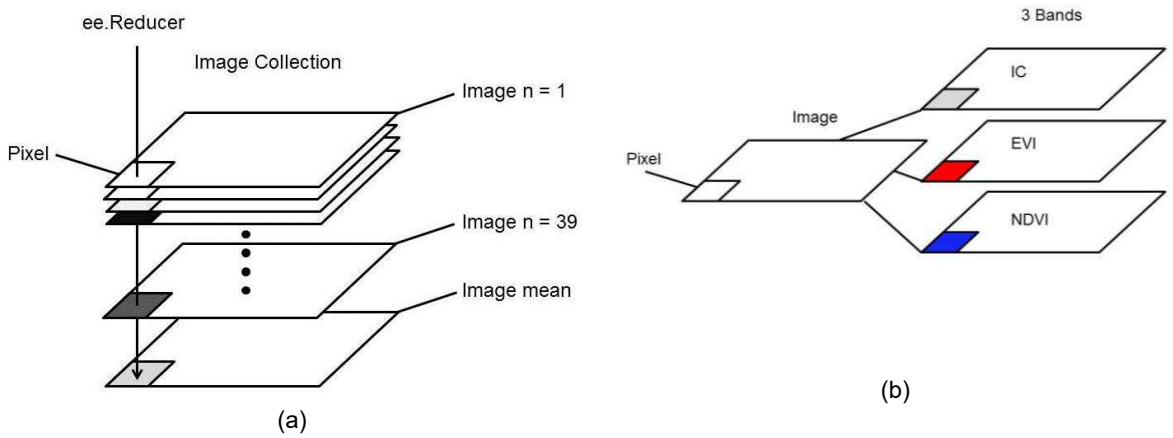


Figure 19. Example of computation of descriptive statistics using the ee.Reducer algorithm across the stack of 39 images (mean in the example) (a). Each image has three bands IC, EVI and NDVI (b).

Because the amount of reflected radiation from each pixel is dependent on IC and this influences VI (Matsushita et al., 2007; Ponzoni et al., 2014), the correlation between EVI~IC and NDVI~IC was analyzed using Pearson correlation to check if changes in IC are related to changes in VI (Zar, 2014). Pearson correlation has shown to be a good descriptor of the linear relationship between IC and VI in previous studies (Galvão et al., 2011; Ponzoni et al., 2014, 2010). The correlation is also computed using the ee.Reducer algorithm and the 39 pairs of values present at each pixel using IC as X and VI as Y (EVI or NDVI). The result for each VI~IC correlation is two images, one containing the values of the Pearson correlation coefficients and the other containing their significance values (p) at the pixel level.

The mean IC, EVI, and NDVI values were calculated for all pixels available at each individual image of the selected images (n=39). This allowed comparing mean values between images at different times. Finally, because IC can be modeled for all the images available (n=397), the mean IC for all of them was calculated to establish the temporal trend. Results were exported to CSV files and analyzed using either R software or EXCEL. Raster images and maps were generated using ArcGIS 10.6.1.

3.3. Socioeconomic approach

3.3.1. Semi-structured interviews and farm typology

20 key informants from the Tropical Agricultural Research and Higher Education Center (CATIE), local government offices of environment and agriculture, forest managers, FONAFIFO officials, municipality staff, and NGO with experience in the PES program were interviewed. These informants, in the first instance, helped to settle concepts and get a first insight into the forestry sector in the area and the functioning of the PES program. These informants also provided a first initial list of farmers participating in the PES program. Once the farm was visited, information and contacts for other farmers were collected. Also, once in the farm location, neighboring farmers were asked if they would like to be interviewed. Farms were selected trying to cover a large variability of managements and land uses, trying to be representative of the landscape observed in the area.

Between the years 2016-2017, a total of 67 semi-structured interviews were collected in the study area following socio-anthropological and qualitative methods (Sibelet et al., 2019, 2013). The inclusion of farms and interviews stopped following the saturation approach, this is when the interview of an additional farm did not provide new information.

During the interviews some topics were raised with the farmers trying to get some insight into the context of farm strategies and the PES program (Arriagada et al., 2009; Fiorini et al., 2020; Sibelet et al., 2017), main topics were:

- A short introduction of the farmers' history, lifestyle, and education
- The main source of income for the farmer, and, if coming from the management of the farm, the land use responsible for that
- The scale of the activities or production carried out on the farm and the technology employed
- Factors influencing farm management and trends in the area
- Farmers’ perception of trees and forest in the landscape and other environmental concerns in the area, such as forest legislation or potential forest uses
- Farmers’ perception of the PES program, knowledge of the program, strengths, weaknesses, and their decisions to participate or not

With all the information collected, a typology of farms was built, grouping farms into types that shared similar management, economic strategies, and farmer profiles as well as similar biophysical characteristics. Some of the responses of the respondents are showed to demonstrate the usefulness of the qualitative information obtained. The qualitative evidence should be evaluated together with the quantitative analysis.

4. Results

4.1. Vegetation risk loss

4.1.1. Random Forests to model vegetation risk loss

The OOB error rate for the prediction of the vegetation loss risk class produced by the RF model was 12.80% and 24.70% for the no-loss risk class. The mean estimate of OOB error rate was 18.71% (Figure 20). It was possible to decrease the OOB error rate to the minimum value using 500 trees (trees), 9 randomly drawn variables (mtry), and 2 as node size (n). Looking at figure 20, the OOB error rate decreases quickly when it reaches 100 trees and seems to stabilize when 200-300 trees are employed, which is interesting to reduce the computation power of the model in future trials.

The RF model produced a classified map showing areas with a predicted risk of vegetation loss (class=1) and no risk of vegetation loss (class=0). The model predicts that 42,836.49 hectares of vegetation cover > 1% in the year 2000 (28.51%) are at risk of being lost. This figure excludes the already lost vegetated areas during the years 2000-2018 used as historical information (Figure 21).

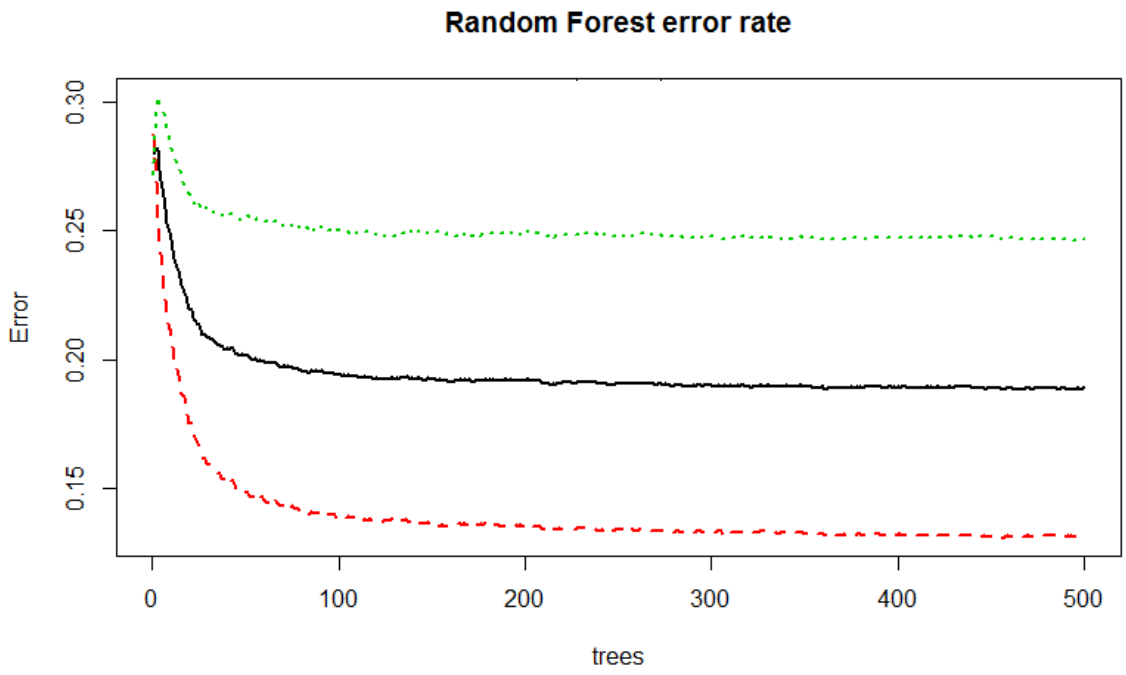


Figure 20. Mean error of the model in black (18.71%), the error of predicting vegetation loss risk in red (12,80%), and the error of predicting non-vegetation loss in green (24.70%).

Additionally, for each pixel in the image, the RF model calculates the probability that the pixel was classified in the correct class. In the case of binary classification, RF looks for the class which wins the majority voting for that data pixel and RF predicts that class for that point. Probability values move within the range 0-1, if the value is 0, it means that the model is very confident about classifying the pixel in class 0 (no vegetation loss), if the value is 1 the model is confident in classifying the pixel in class 1 (vegetation loss). Any value between 0 and 1 represents the remaining possible probabilities that the pixel is classified in any of the two classes. The cutoff value of 0.5 represents an equal chance of being classified in any of the two classes by the model or that the model recognizes equally large numbers of votes in favor and against classifying it to any of the two classes. This probability can also be transformed into a vegetation loss risk scale that shows the likelihood that vegetation is lost or not within a pixel. Using the intervals [0-0.1]; [0.1-0.25]; [0.25-0.5]; [0.5-0.75]; [0.75-0.9];[0-9,1], the proportion of area at risk of vegetation loss is from very low to very high 39.48%, 13.72%, 18.29%, 14.97%, 8.62% and 4.92% respectively (Figure 22).

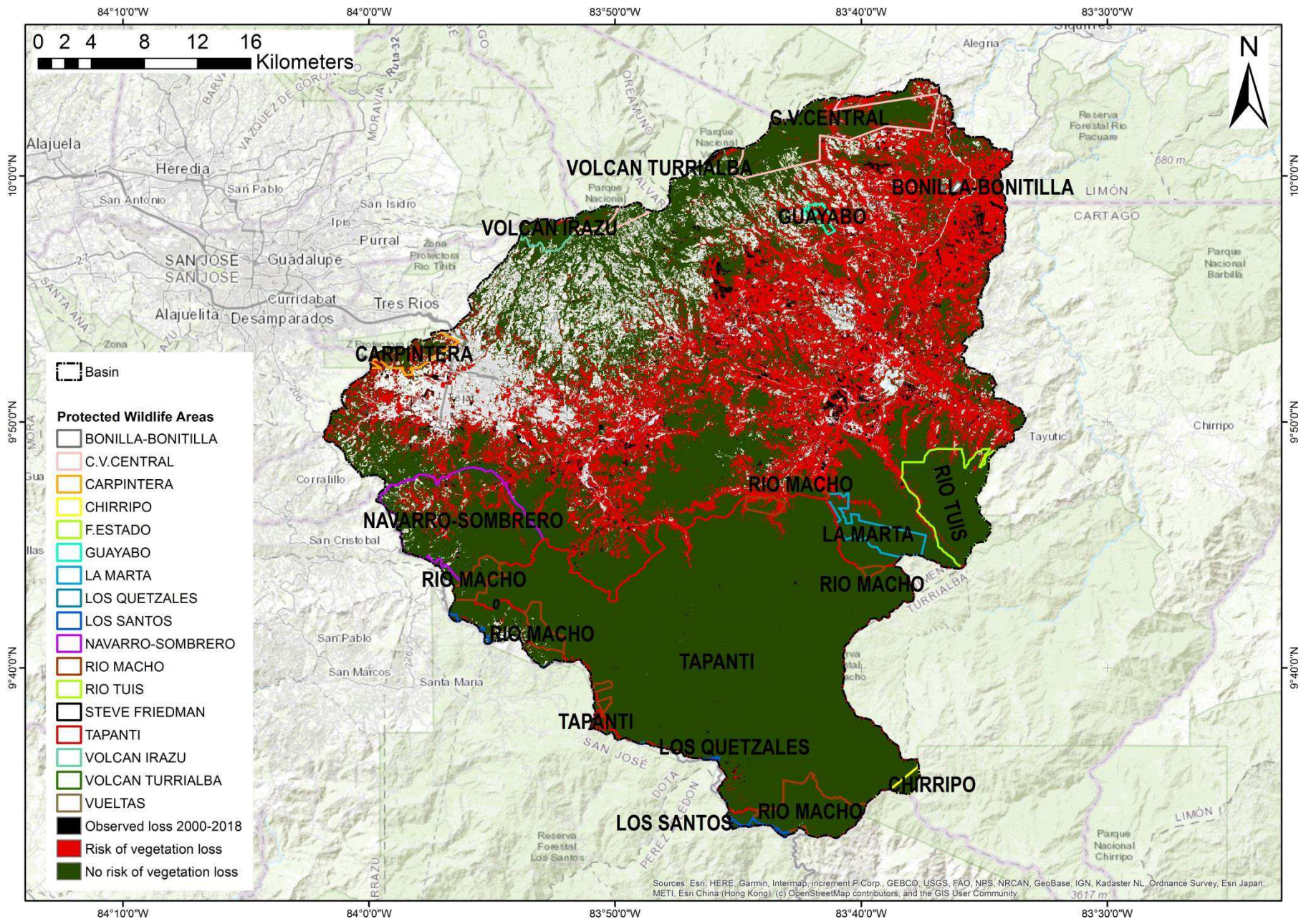


Figure 21. Predicted vegetation loss map in the study area. Black areas represent historically lost areas. Red color represent predicted loss areas and green predicted no loss areas. The map of Protected Wildlife Areas is overlaid.

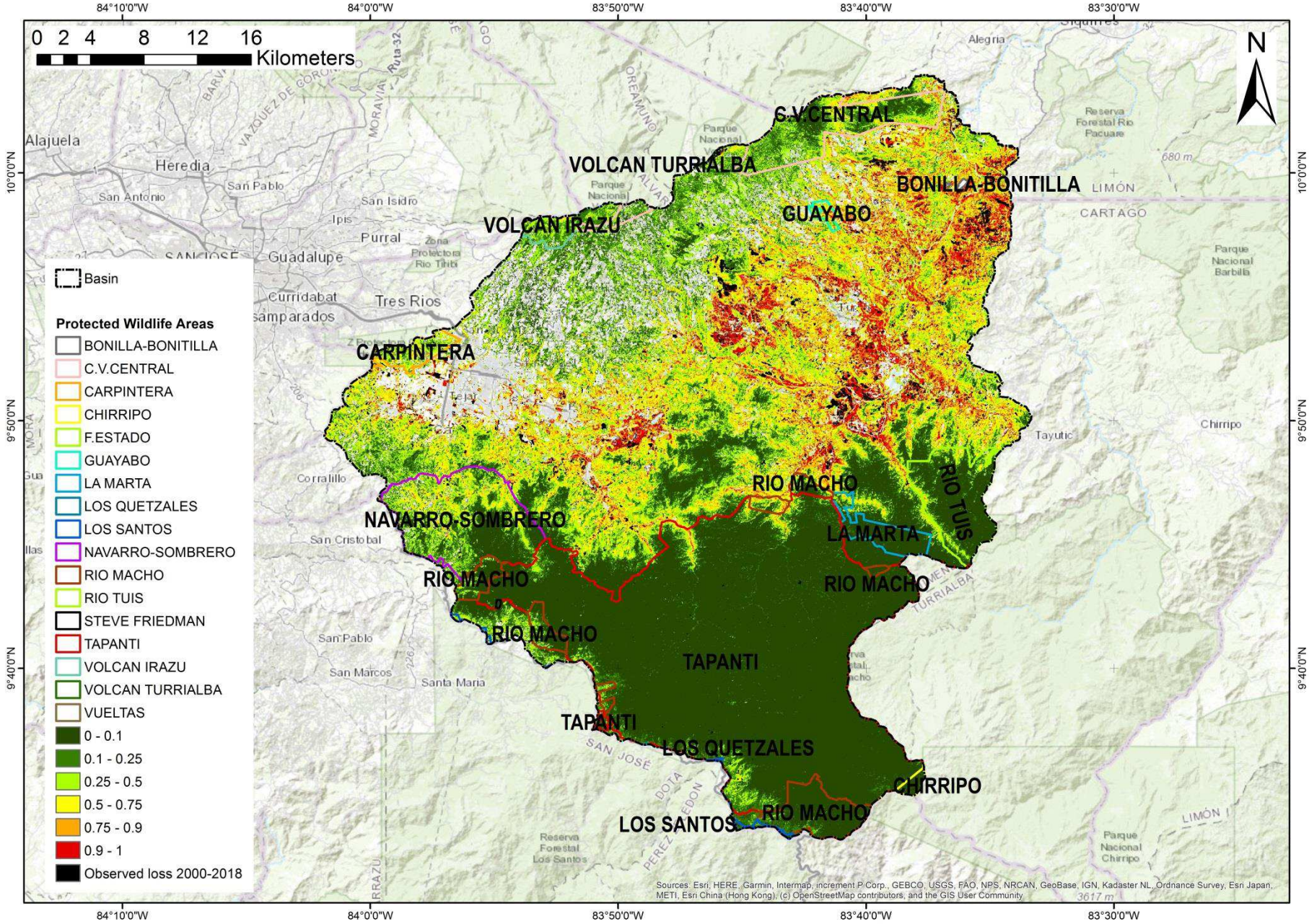


Figure 22. Probability map for predicted vegetation loss risk map in the study area. Black areas represent historically lost areas. Red color represent areas with the highest probability (0.9-1) of being correctly classified as vegetation loss. Dark green represent the highest probability (0-0.1) of being correctly classified as no loss vegetation.

4.1.2. Validation of the model

The accuracy and reliability of the binary RF classification were evaluated by estimating its Area Under the Curve (AUC) of Receiver Operating Characteristics (ROC). With an AUC of 0.89 the reliability of the model was good (Figure 23), indicating that the model is accurate at predicting vegetation loss, similar values have been found in ecological studies (Dlamini, 2016; Drew et al., 2010; Humphries et al., 2018; Saha et al., 2020).

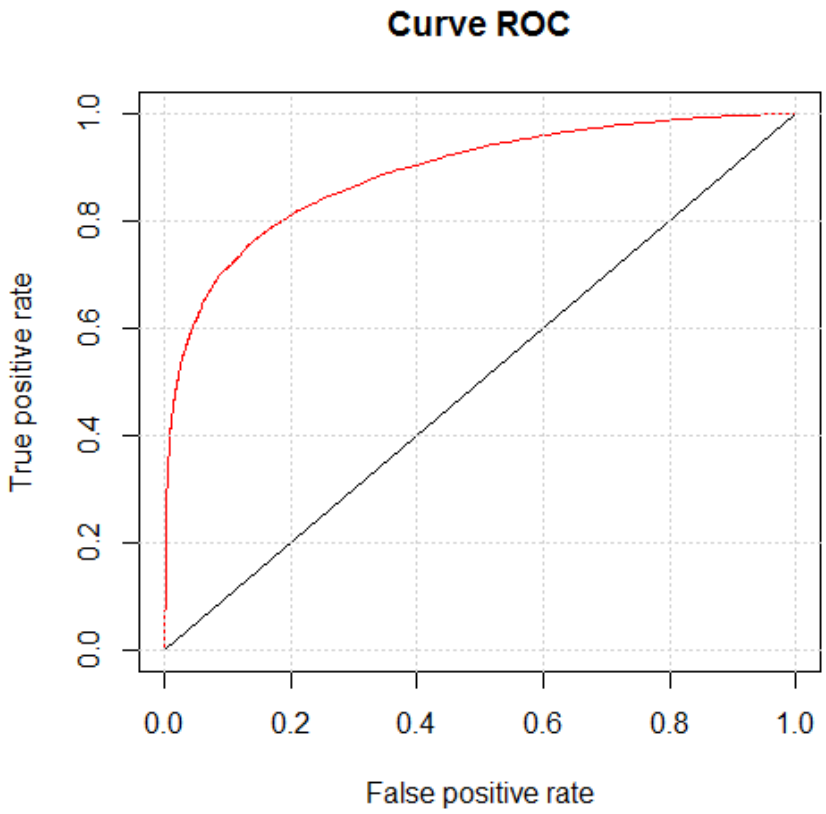


Figure 23. AUC value of 0.89 after independent validation.

4.1.3. Importance of the predictor variables

The predictive impact of each predictor variable was calculated for the model. Accessibility and topography factors such as distance to main roads, elevation, slope, and aspect had the greatest impact in decreasing order for both MDA and MDI (Figure 24). Among the different indices, LSWI and NDTI were considered relatively more important based on MDA, and the remaining indices did not show large differences between them.

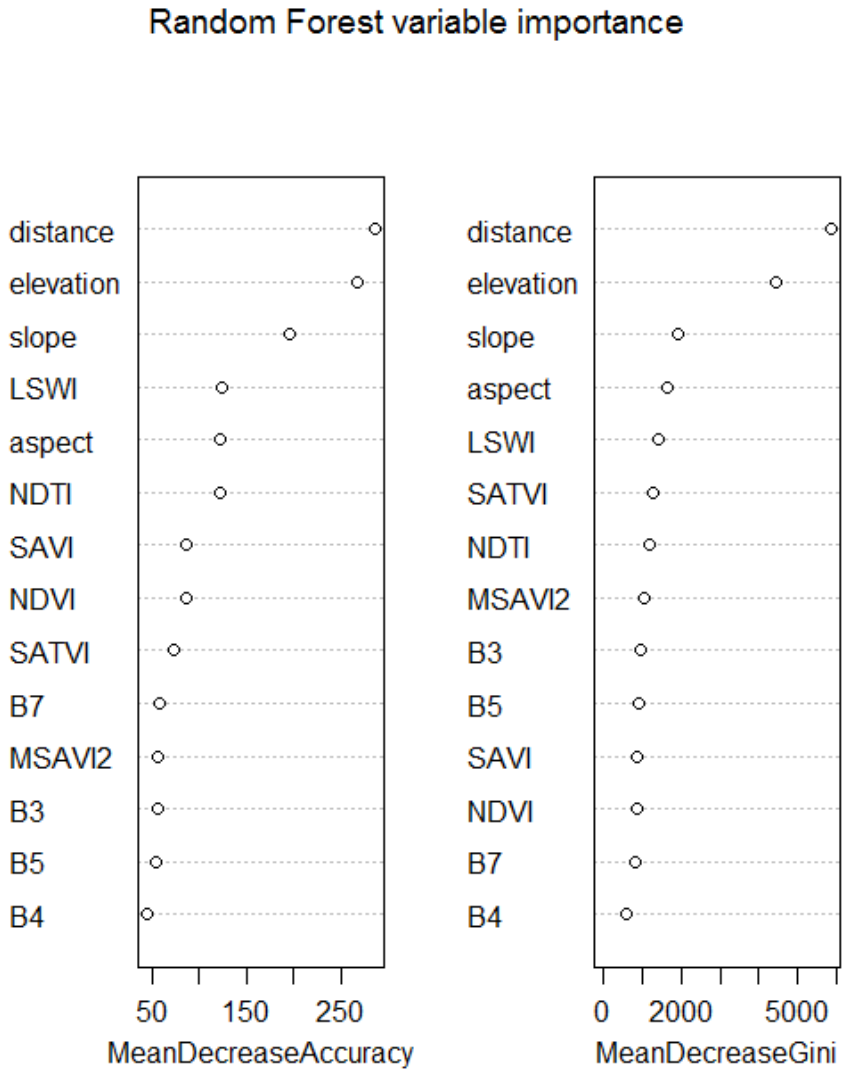


Figure 24. Predictive ability of each predictor variable in explaining the presence of the absence of vegetation loss in the model. Higher values represent higher importance.

Partial dependency plots are illustrated for the four most important variables following importance ranked by MDA. The plots illustrate the marginal effect measured from 0 (low) to 1 (high), that the variables distance, elevation, slope, and LSWI have on the accuracy of the model to predict vegetation loss (Figure 25). The effect for all of them is stronger with low values, except for LSWI. It continuously decreases until reaching its minimum effect as the values increase and then increases again when it reaches higher values. In the case of LSWI, the effect is stronger with higher values.

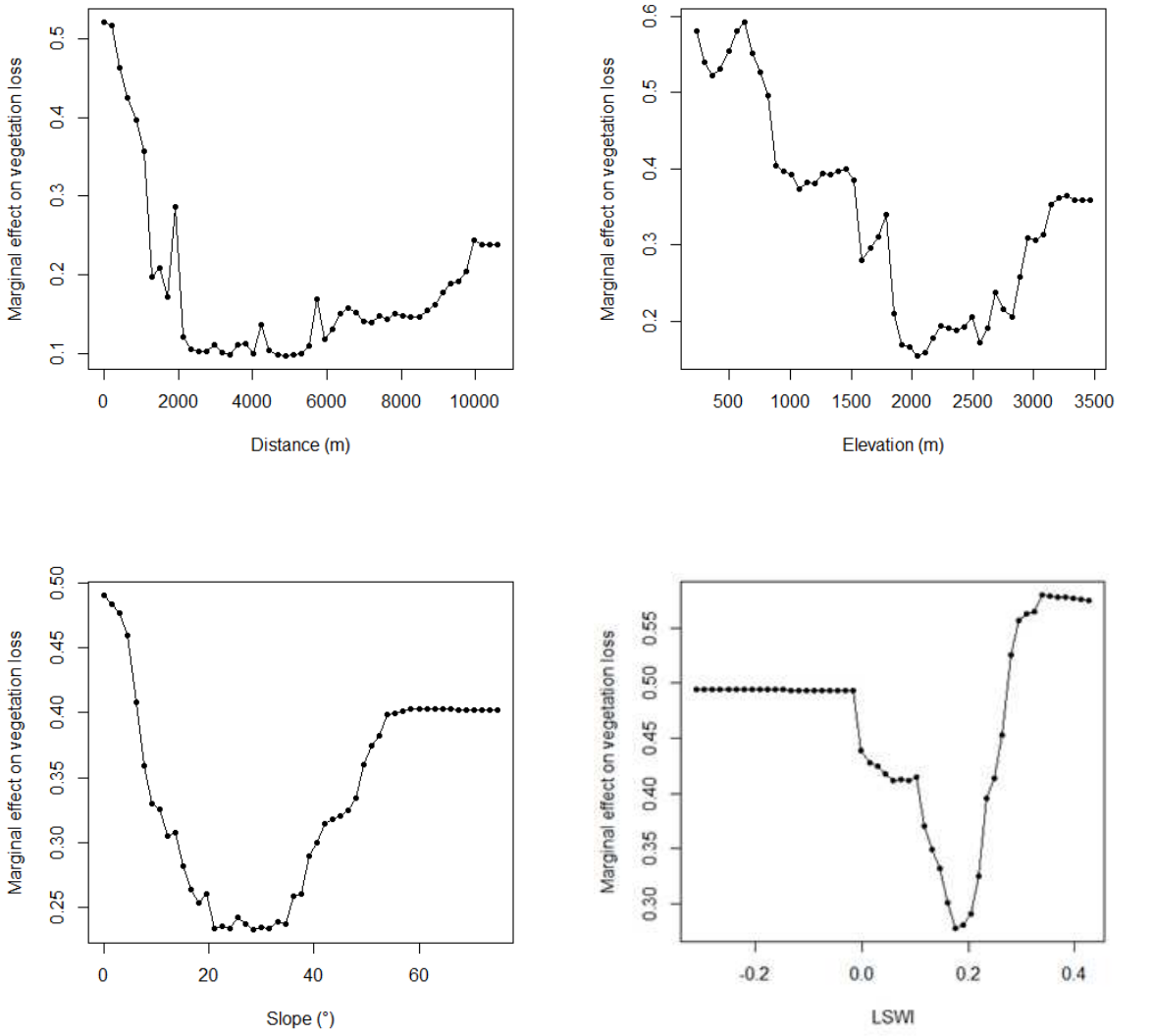


Figure 25. Partial dependency plots for Distance to main roads, elevation, slope, and LSWI. The vertical axis shows the marginal effect of the predictor variable on the accuracy of the model as units of the variable change in the horizontal axis.

4.1.4. Historical and predicted vegetation loss risk between protected wildlife areas and unprotected areas

Historical vegetation loss during the years 2000-2018 for each area considered was well below 5%, except for Volcán Turrialba that lost 9.38% of its vegetated cover (Figure 26). In terms of total area size, the majority of vegetation loss occurred outside the PWA, in the basin, which lost an absolute number of 4,163.22 ha. In comparison, the combination of all observed losses for all PWA was only 213.57 hectares. The lowest loss occurred in Finca Estado, with an absolute value of 0.09 hectares. Several areas such as La Marta, Los Quetzales, Steve Friedman, and Las Vueltas remained unaffected during the historical period.

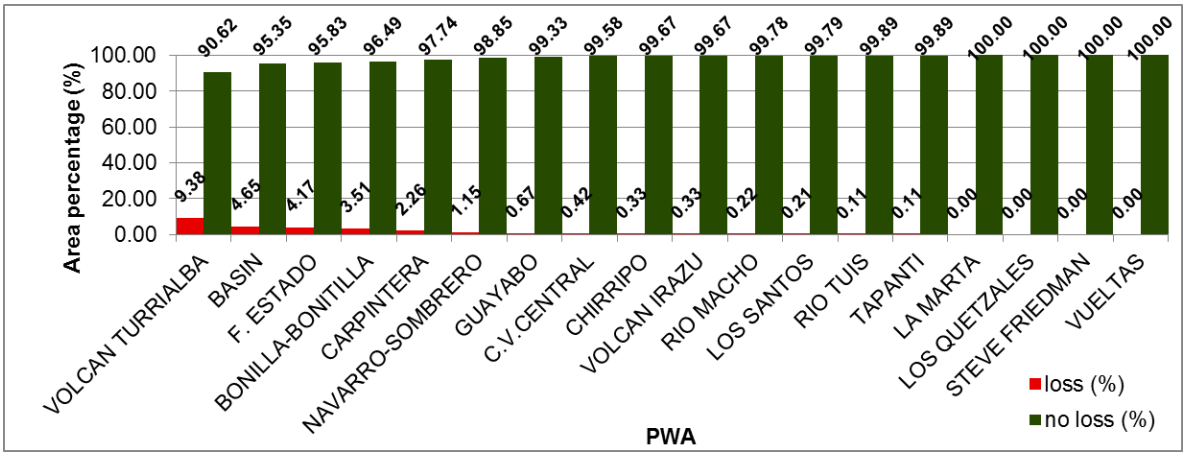


Figure 26. Historical observed vegetation loss and no loss areas inside the protected wildlife areas in the study area. Bars are ordered from left to right in decreasing order of area lost.

The analysis of the predicted vegetation loss from the year 2000 onwards, shows an increased percentage of vegetation loss for all areas except Chirripó, Steve Friedman, and Vueltas that would remain unaffected. All the figures shown in the predictions exclude already lost areas. Following the model, 28.51% of the total area or 42,836.49 hectares in the study site are at risk of being lost, Again, the most affected area in terms of area size is the Basin, therefore outside the PWA. The model predicts that 47.27% or 40,215.60 hectares of vegetation from the year 2000 onwards will be lost inside the Basin, which is 10 times higher than the historical loss observed. The next largest loss in terms of area is predicted to occur in Navarro-Sombrero with 921.96 ha or 15.22% of its size, which compared with the predictions in the Basin is quite low. In terms of percentage, the Bonilla-

Bonitilla wetland is the most affected PWA with a predicted loss of 96.36% of its area, although it only accounts for 9.9 ha in size (Figure 27).

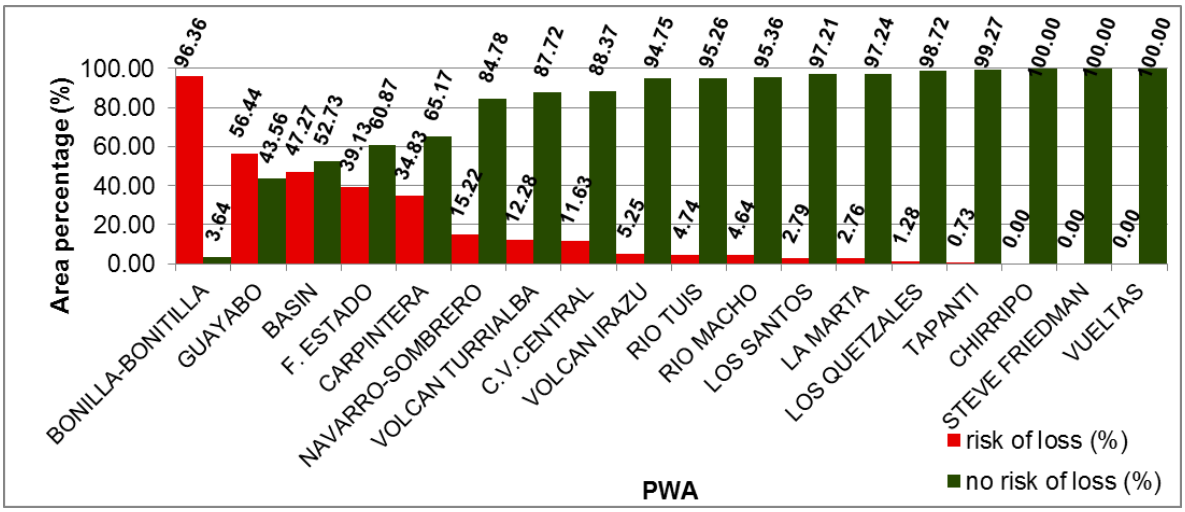


Figure 27. Predicted vegetation loss and no loss areas by the model inside the protected wildlife areas in the study area. Bars are ordered from left to right in decreasing order of area lost.

Further analysis of the predicted loss was made using the probability intervals selected from the RF model (Figure 28). This provides a higher detail in considering the likelihood of vegetation loss in an area. Considering the probability intervals that we transformed into a vegetation loss risk scale, only 8.55% or 7,300.35 hectares of the Basin area would be at very high risk [0.9-1] of being lost. Bonilla-Bonitilla and Guayabo would be the following areas with the largest proportion of area in the very high-risk interval with 24.55% (2.43 hectares) and 8.22% (17.46 hectares) of their respective areas at very high-risk. The rest of the areas all show very high-risk intervals well below 0.5%.

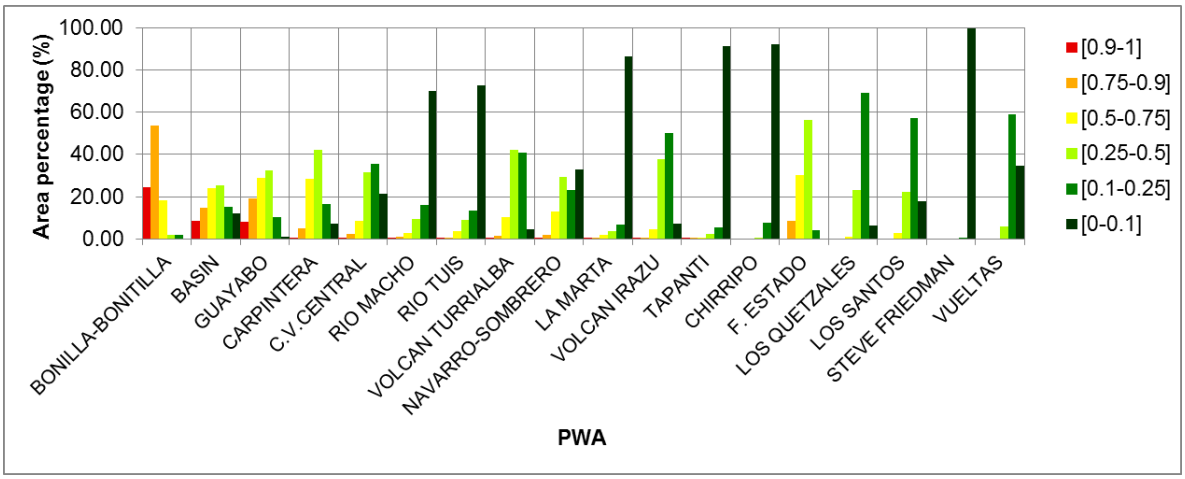


Figure 28. Intervals of risk probability for predicted vegetation loss and no loss areas by the model inside the protected wildlife areas in the study area. Bars are ordered from left to right in decreasing order of area at higher risk.

More than 90% of the vegetation located outside the protected areas is at risk of being lost, following the risk intervals [0.5-0.75], [0.75-0.9] and [0.9-1] (90.79%, 96.51%, and 98.68% respectively) (Table 4 and Figure 29).

Table 4. Risk of vegetation loss in the study site by probability intervals in hectares and as a proportion of protected and unprotected areas

Vegetation loss risk probability	Total area (ha)	Protected Wildlife Areas (%)	Non-Protected Wildlife Areas-Basin (%)
0.9-1	7,398.36	1.32	98.68
0.75-0.9	12,949.2	3.49	96.51
0.5-0.75	22,488.93	9.21	90.79
0.25-0.5	27,475.47	21.35	78.65
0.1-0.25	20,612.43	36.58	63.42
0-0.1	59,326.02	82.25	17.75

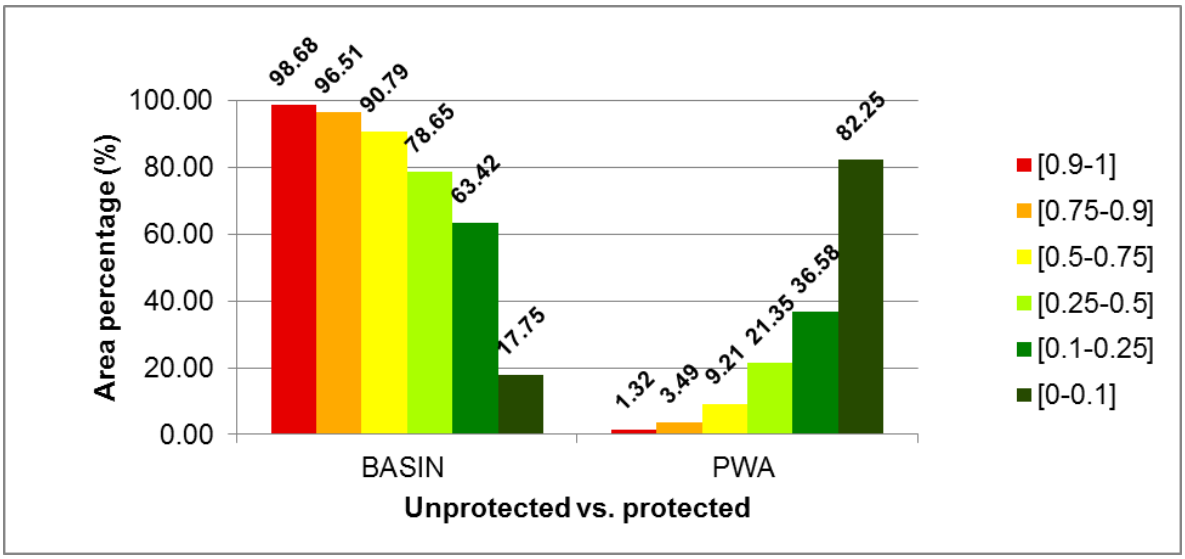


Figure 29. Proportions of area under different risk probabilities between protected and unprotected areas.

4.2. Farms dynamics

4.2.1. Typology of farms

A farm typology was elaborated with different management strategies (Table 5 and figure 30). Discriminant factors were used to separate types. Factors in order of importance were: principal land use within the farm responsible for the principal income/use, the scale of the activity, technological level, and level of education or social capital. Based on these factors 6 types were developed. Type A is traditional farming with agriculture as the main activity. Type B includes farmers that also have agriculture as the main activity, but they have a higher educational level that allowed them to diversify their products and be more efficient in the use of the resources and more environmentally friendly. Type C corresponds to larger areas and productions, usually belonging to cooperatives. Sugar cane and coffee are the main crops and small to medium processing plants and infrastructures are located within the farms. Type D includes owners that dedicate most of the land area or infrastructures to animal breeding, from dairy cattle in pastures, pigs in stables, or trouts in small water reservoirs. Type E are farms where the forest area provides the main income from tourism, like hostels or private conservation reserves. Type

F is represented by large farms with a high proportion of forest. The property is used for leisure, without any productive activity taking place in the property.

Table 5. Typology for the different types of farms

Type and number of farms of each type within the sample (n=67)	Characteristics of farms and farmers' strategy
A Traditional agriculture (10)	-Small to medium area and production (less than 20 hectares), self-consumption, or local markets -Diversified production with traditional crops, potato, carrot -Basic technological level -Low educational level
B Innovative agriculture (8)	-Small to medium area and production (less than 10 hectares), self-consumption or local market -Diversified production with the inclusion of less traditional crops, strawberry, avocado, biological agriculture -More advanced technological level, compost, earthworms, water reservoirs, own machinery -Higher educational level
C Technified agriculture (10)	-Medium to large area and production (more than 20 hectares), local markets or cooperatives -Little diversification, coffee, sugar cane or production of agricultural crops -Use of machinery -Small processing plants, crop processing, crop packaging, warehouses
D Animal breeding (14)	-Dairy cattle production, pigs, trout -Professional technical team -Most of them belong to cooperatives i.e., to sell the milk)
E Tourism/research (11)	-From small familiar businesses (restaurants, hostels) to private conservation reserves and experimental University sites -They exploit the landscape ("forest") as source of income
F Leisure (15)	-The owners usually live in the property or close -The main income does not come from the farm -Involved in local organizations, local initiatives, in relation to other stakeholders and institutions -Retired high-level workers, technicians, politicians, lawyers

Farm size ranged from 0.45 to 1,563 hectares with an average size of 143 hectares. The majority of the agricultural lands were located under the slopes of Irazú volcano, which although at high elevation and steep slopes provide a very deep soil of volcanic ash, rich in mineral nutrients. This area is also relatively less cloudy throughout the year and is

subject to more insolation. Conversely, farms with less area under agricultural or productive land have more forest area and are located on the opposite side of the basin, close to Tapantí National Park. This area is very cloudy and fog is frequent throughout the day, especially in the evening.

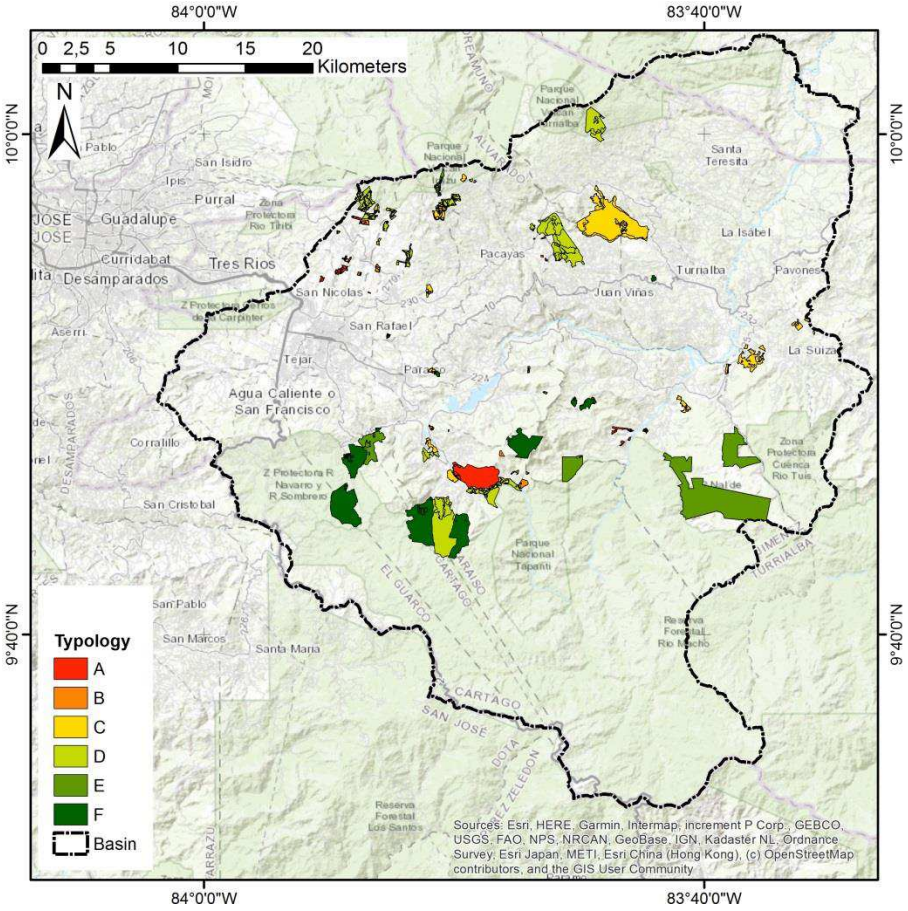


Figure 30. Location of the farms visited categorized using the typology.

4.2.2. Land use share, forest and opportunity cost

The share of different land uses inside the farms varies across the different types in the typology. Land uses within the farms were named agriculture (agr); sugar cane (can); coffee (cof); forest (for); infrastructure (inf); pasture (pa) and pasture with trees (pat). The areas of all land use for all farms within each type were added and were shown as an individual percentage of the total land area per type (Figure 31). Concerning forest land use, farms in type C have the lowest proportion and type E the highest. Sugar cane or coffee occupy large extensions within the farm because type C is the closest to industrial

production. There is little space to allow the forest to grow as that would incur a high opportunity cost. Moreover, these farms are frequently on gentle slopes easy to work and access. Then from types A, B, and D as the activities in the farms diversify, and as the topography limits crops, it is possible to find larger areas with forest. They are frequently found between different land-use areas or protecting steep slopes close to small rivers crossing or dividing the properties. Forested areas are also found in the most inaccessible parts of the farm, with lower opportunity costs.

Types E and F exhibit the highest share in forest area. The whole farm is covered by forest with small parts dedicated to other uses or where buildings are placed. The largest farms with forest are placed at higher elevations compared with the rest of the types and accessibility due to the terrain is complicated. The alternative to placing productive crops in these shallow and steep soils is very unlikely and therefore opportunity cost is the lowest. An indicator of the location of the forests is how respondents use the term "montaña" -mountain in English- when referring to the forested areas.

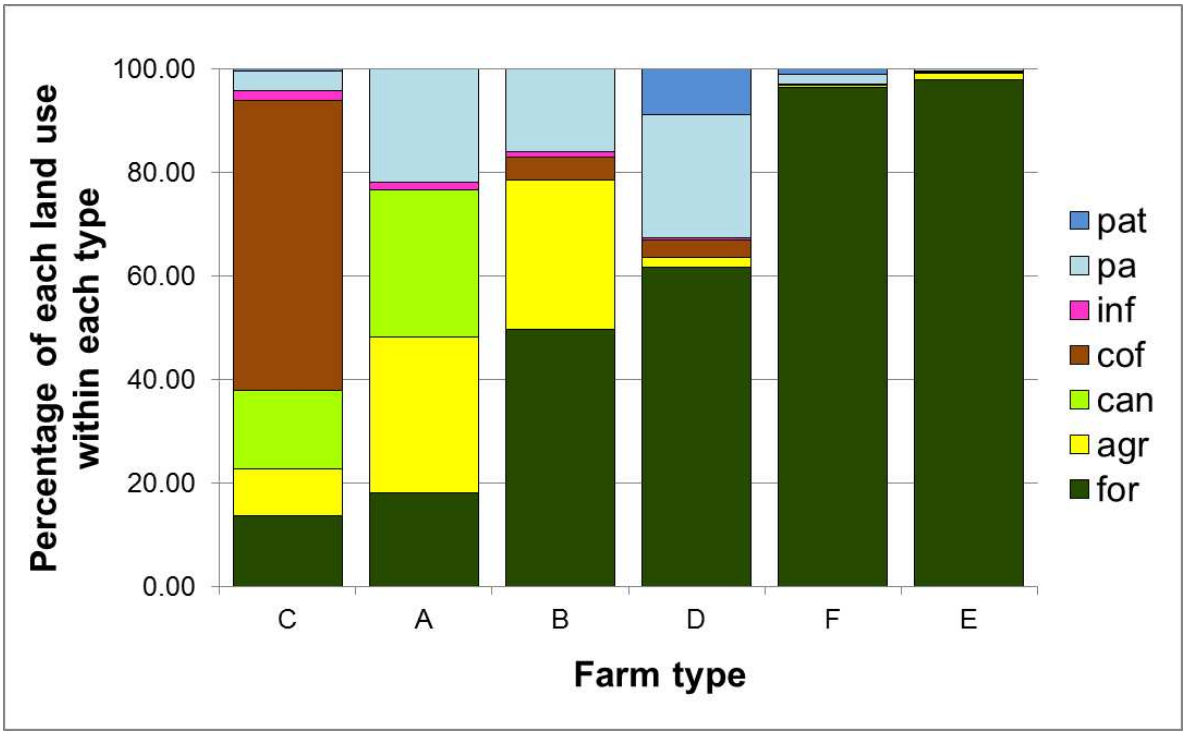


Figure 31. In order of increasing share of forest area. The plot excludes one farm with 434 hectares of forest in type A and 3 small farms (0.81 to 3.15 hectares of total area) without forest in type F which are not representative in terms of percentage of forest area from the total area.

When the share of forest area was analyzed for all the farms, a positive linear relationship was found between the share of forest area inside the farm and the total farm area (Figure 32). When the factor of participating or not participating in the PES program was added, there seems to be a combination of forest share and total farm area below which is not frequent to find farms participating in the PES program. Excluding a farm receiving payments for agroforestry (2.79 hectares), the rest are all included in the forest protection modality. The farm with the minimum forest area participating in the program had 15.93 hectares of forest out of 18.45 hectares (86.34%). The farm with the largest share of forest area included in the PES program has 482.94 hectares of forest out of 602.46 hectares in total (80.16%).

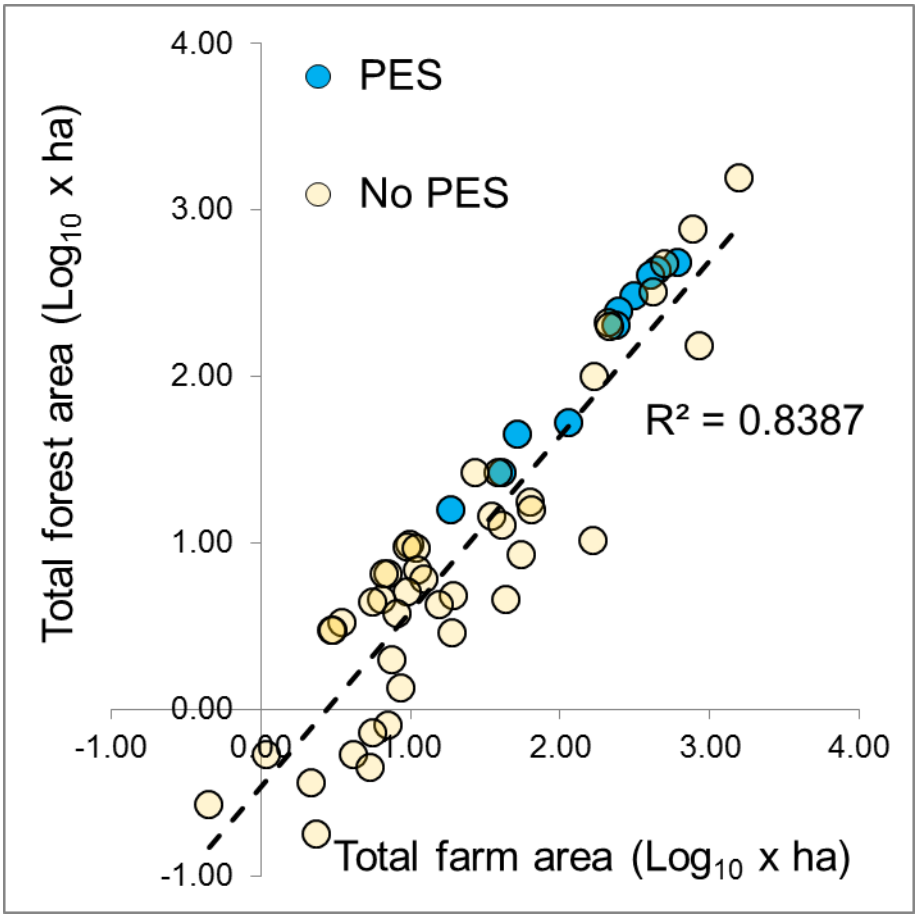


Figure 32. Scatterplot showing the areas of all farms visited and the proportion of forested area in each farm. The smallest farm with PES, receives the modality of Agroforestry System, which entails the plantation of trees to create an agroforestry system, the size of forest in that farm is 2.79 ha. The next largest one has 15.93 ha of forest. Areas have been transformed using Log_{10} for ease of visualization.

Productivity and accessibility are fundamental variables to explain this. Very productive soils such as those dedicated to agriculture can give several productions per year, are easy to work, and are deep in mineral volcanic ash nutrients. Because of the high productivity, the farms can be divided into small parts, and, also when inherited, the farms are split between family members. For example; 7,000 square meters (0.7 hectares) for agriculture is a common unit used in the study site to measure the size of a farm, and frequently enough to generate the income of a farmer. Besides, they are well connected with main and secondary roads, close to urban centers so that transport costs are diminished. Conversely, large farms with larger forested areas are located in more remote areas, and soil productivity is very low. There is no production associated directly with the trees and most of these farmers declared that the only alternative to the presence of forest could be the development of tourism, related to guided tours or bird observation, which requires the permanence of forest. Additional factors that affect farm size involving economic and legal aspects. Farms dedicated to agriculture are expensive, they are productive and interesting from a short-term business perspective. The accessibility and low tree coverage make it very easy to measure them and draw property maps. Respondents in the area described a relatively active market for agricultural lands, also from people interested to build their own house. On the other hand, large forest farms, are less interesting from a short-term business perspective, but can be interesting as an investment asset. Their soils cannot be compared in terms of production and the land market is not so active. Based on some respondent´s experiences, accessibility is a factor affecting how the properties are delimited. Densely forested farms with irregular terrain are more difficult to delimitate, it can take several days to walk its perimeter. Usually, maps are not clear or updated, which often leads to disputes between owners and hinders participation in the PES program.

“It is hard to believe that our neighbor with 500 ha steals 90 ha from us” [Respondent 1]

The relation between, accessibility, productivity, and size can be observed using the cadastral maps available at the Costarican National Geographic Institute (https://www.snitcr.go.cr/). Figure 33 shows available cadastral maps for the Municipality of Orosí and its surroundings. Accessible properties dedicated to agriculture and coffee are small, with clear boundaries. As forest becomes predominant and terrain access more difficult, farm size increases greatly, and straight lines can be observed between large

farms which also casts doubt on the precision of its limit measurements. Delineation of farm boundaries is at the center of most of the disputes between owners, and if the legal property is not clear, the farm is prevented to enter the PES program.

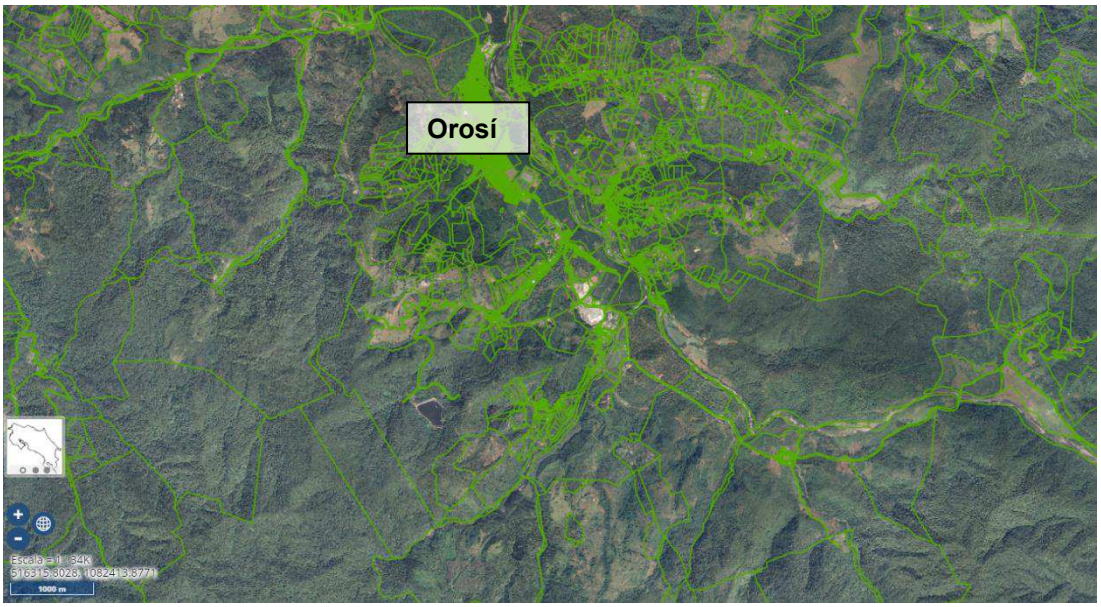


Figure 33. Cadastral map around the municipality of Orosí, close to the Tapantí National Park. Farms located far from the urban area in less accessible areas are larger in size and in forest area.

4.2.3. Historical and predicted forest loss risk inside delineated farms

The results of historical vegetation loss displayed in this section, correspond to vegetation loss occurring between 2000-2017, or right after, in 2018, over observed forested areas in 2017. Because the forest area in the farms was delineated and present in 2017, there are two ways to interpret vegetation loss: 1) The vegetation loss occurred any year before 2017 and the delineated forest in 2017 is the result of regrowth, or 2) The vegetation loss occurred in 2018 and then the data is showing forest loss from the delineated forest in 2017. In general, for forested areas in all farm types, vegetation loss was very low during the historical period (Figure 34). This means that the forest observed in 2017 has been permanently present over more than 94% of its area for all farm types throughout the period studied and regardless of having participated or not in the PES program. The vegetation loss was below 6% of the total forest area for all types. In general, farms not participating in the PES program showed higher vegetation losses than those participating, except for types E and B that showed the same losses.

In terms of total area size, the majority of vegetation loss occurred for type D not participating in the PES program, and based on the field interviews and visits corresponds to selective logging of 29.88 hectares of tree plantations converted to pasture land between the years 2015-2017 in a particular farm. The loss observed in type A in the same no PES group is represented by 0.36 hectares of forest lost in the border of sugar cane plantations. Type C corresponds to 3.00 hectares of forest lost surrounding coffee plantations, although it could correspond to loss of coffee shrubs as well. Loss in type F also corresponds to 15.93 hectares of planted forest logged in the year 2018, and the remaining losses in type E correspond to small areas (< 1 hectare) of trees removed to expand agricultural areas.

For farms that participated in the PES program, the highest loss is attributable to a farm in type A, where 3.42 hectares of planted trees were logged and transformed to agricultural land use between 2018-2019. Forest loss in type F corresponds to the logging of 1.17 hectares of scattered trees to extend and improve the condition of the roads inside a farm dedicated to leisure. Losses in types F and D are the result of the clearing of 1.17 hectares of vegetation prior to 2009 and the subsequent plantation of new trees in the first case, and the clearing of 1.26 hectares between the years 2015-2016 of planted trees in a particular farm in the second case. Forest losses observed in type E correspond to a particular farm clearing 0.36 hectares of forest for the expansion of pasture and the plantations of fruit trees between the years 2001-2017.

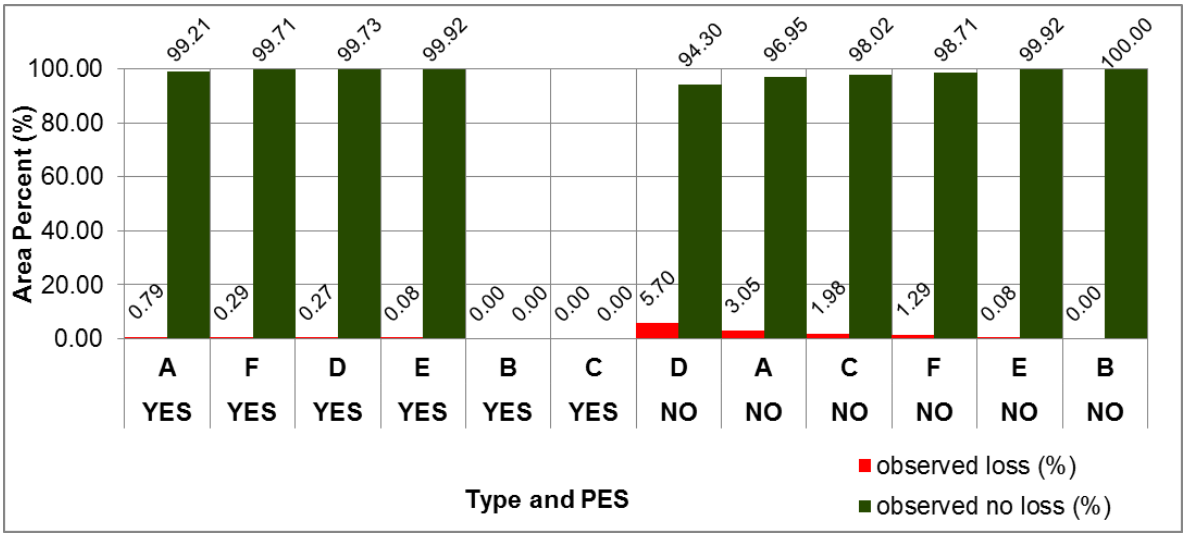


Figure 34. Historical observed forest loss and no loss areas inside the farms in the study area. Bars are grouped by type and participating or not in the PES program. Bars are ordered from left to right in decreasing order of area lost.

Excluding vegetation areas lost in the historical period, the model predicts a significant potential increase in forest loss for all types except for types B and C that participated in the PES program (Figure 35). Types D and C not participating in the PES program show the highest predicted vegetation losses. Type D corresponds mainly to 2 farms. The first one would lose 205.29 hectares of forest, mainly surrounding productive areas and part of this forest would correspond to planted forests with future logging possibilities. When the highest probability of the model is analyzed (0.9-1), the area is reduced only to 12.87 hectares. The second case for type D corresponds to another single farm that would lose about 25.02 hectares of forested areas, mainly close to rivers in flat areas and bordering with pasture and agricultural areas. When the highest probability is taken into account, only 3.15 hectares would be lost. Most of the forest area predicted to be lost in type C corresponds to the borders of a farm where shadowed coffee is the main productive use. The model predicts that up to 54.36 hectares of forest could be lost, and when the highest probability is analyzed, only 7.29 hectares would be affected. Based on the location of the areas at more risk, from an economic perspective, it could be interesting to remove trees to increase coffee production, but also deforestation could be due to landslides caused by the river bordering the farm instead of human causes.

Another prediction of forest loss is for another farm for type C, where 9.18 hectares of forest divided into several forest spots could be lost. The different spots are isolated, like islands, and surrounded by sugar cane plantations (Figure 36), it could be possible a transformation on these areas if the plantation of sugar cane is feasible, because accessibility is high. For type B, one farm would have 1.71 hectares at risk, but with a probability below 0.9.

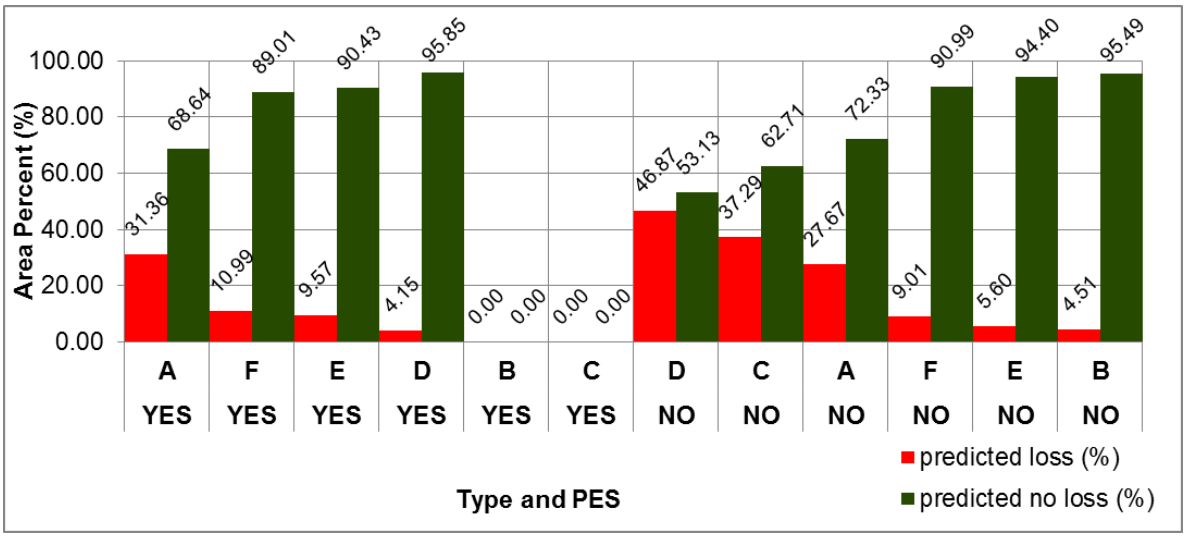


Figure 35. Predicted forest loss and no loss areas by the model the inside farms in the study area. Bars are grouped by type and participating or not in the PES program. Bars are ordered from left to right in decreasing order of area lost.

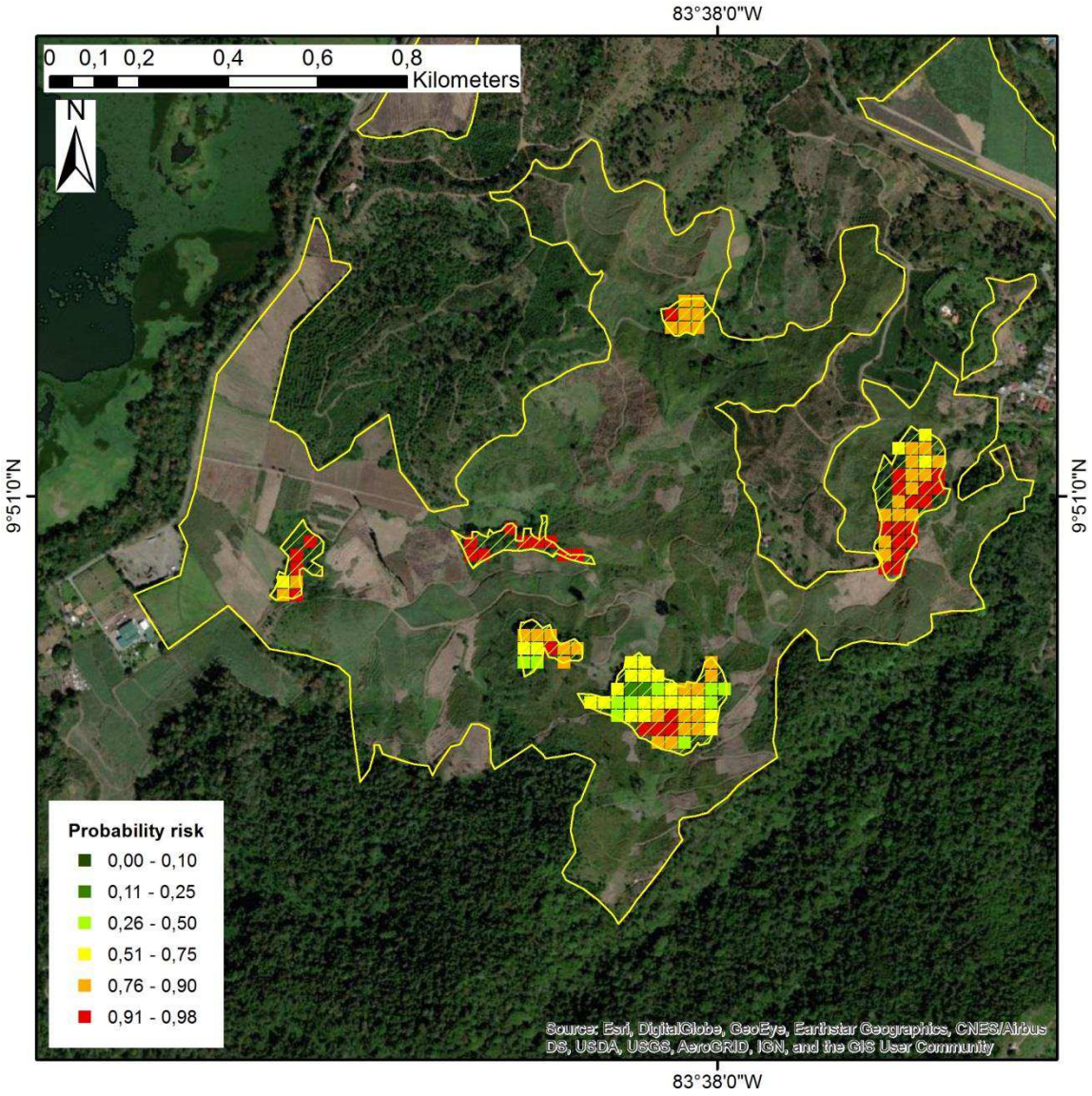


Figure 36. The main land use of this farm is sugar cane. The colored squares (30 x 30-meter pixels) show the predicted probability of forest loss patches remaining inside the farm. Source: Maxar (2020).

For types A, F, and E, farms participating in the PES program have a comparatively higher percentage of forest areas at risk of being lost as a fraction of their total forest area than farms not participating in the PES program. In quantitative terms, 244.26 hectares of forest would be lost in PES farms and 229.86 hectares in no PES farms. However, when only the highest probability interval o 0.9-1 is considered, the area is reduced to 7.2 hectares in PES and 10.98 in no PES respectively. The highest predicted forest loss corresponds to 135.09 hectares of a farm in type A participating in the PES program (Figure 37). Part of this loss corresponds to a remaining tree plantation that can potentially be logged and

because some other areas of the farm can be transformed to agricultural areas or pastures as it has occurred in the past. A large extent of the forest was cleared for agricultural land close to the border with neighboring farms, where the model is predicting the loss and which would indicate an opportunity to develop productive activities. Areas with more risk probability are close to river streams, where part of the loss could be explained by flooding and change of the water course as previously occurred. Selecting the higher range of probability the risk is greatly reduced to only 4.32 hectares.

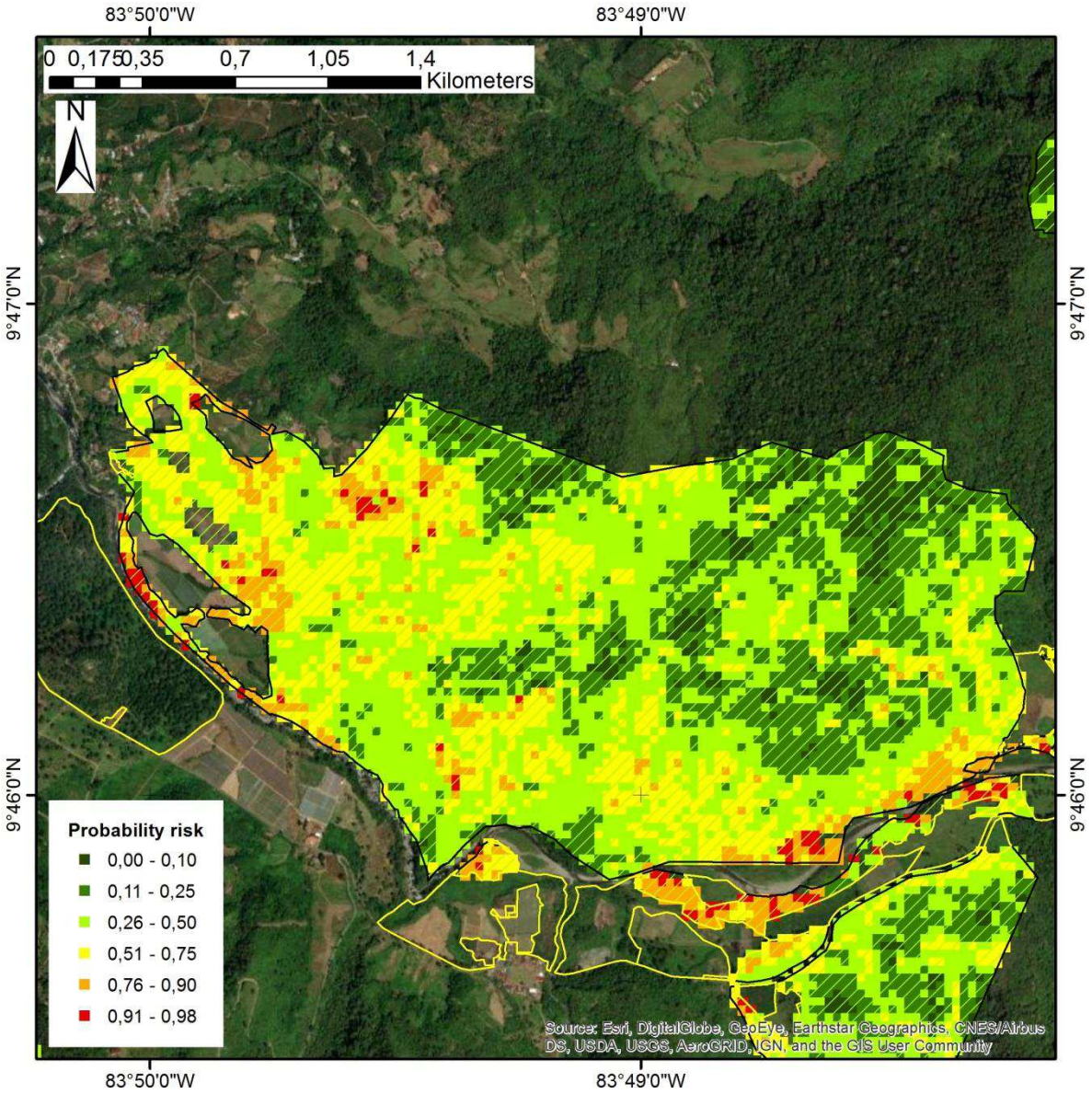


Figure 37. The main land use of this farm is forest. The colored squares (30 x 30-meter pixels) show the predicted probability of forest loss patches remaining inside the farm. Source: Maxar (2018).

Predicted loss for type F mainly corresponds to two farms where 42.21 and 24.93 hectares of forest are at risk of being lost. Predicted loss is close to previously observed loss, for pasture expansion or road improvement, but when the risk interval of 0.9-1 is considered, only 1.08 hectares remain at high risk for only one of the farms. For type E, the risk seems to be related to a potential expansion of agricultural land close to the river stream, or due to the river itself, which can provoke landslides and transport large rocks during storms as it was responded in the interviews. The area at risk passed from 38.61 hectares to 1.71 hectares when the highest probability (0.9-1) was considered.

For type A in the non-PES group, most of the predicted loss corresponds to 3.60 hectares of forest surrounding sugar cane plantations, which serve as farm boundaries and also surrounding buildings. This area could also be cleared from an economic perspective if needed.

Figure 38 shows the percentage of forest in each farm type at risk of being lost, with detailed information of risk probabilities. Clearly, E and F types have lower risks of forest loss, because forest conservation itself is an objective in farm management. Having as principal objectives, tourism, conservation, or leisure, a great loss of forest would be negative for their purposes.

"Tourism brings more benefits than PES. The program pays very little and the paperwork is too tedious" [Respondent 2]

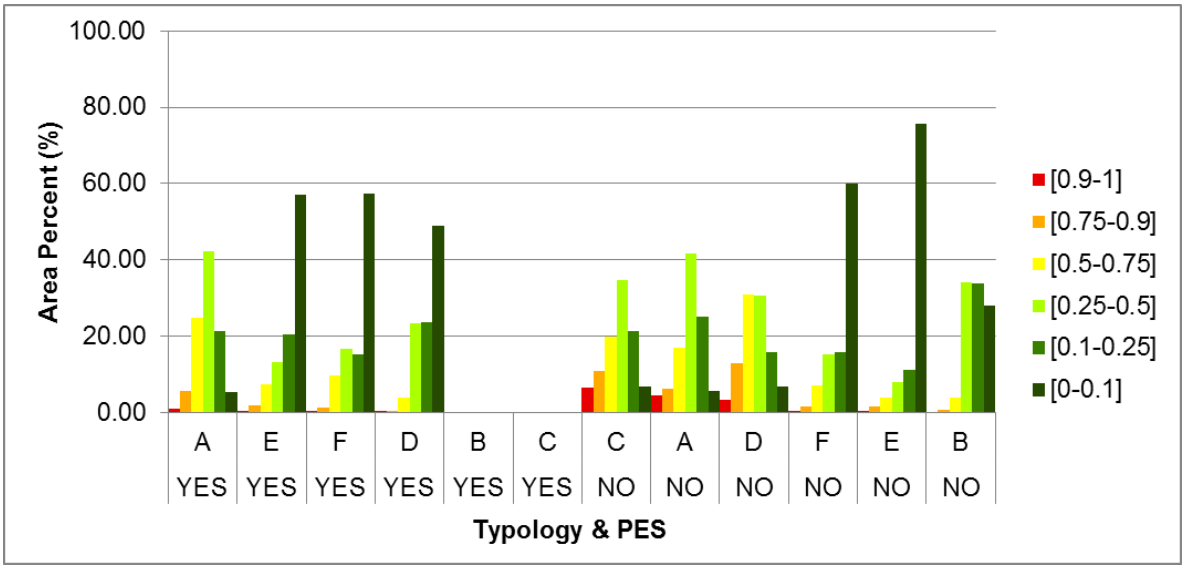


Figure 38. Percentage of forest area at risk of being lost for each type and PES. Letters indicate types. YES= participating in the PES program; NO= not participating in the PES program. Colors indicate the different risk intervals.

Since the location of forest concerning productive land uses, seems to be an important factor defining the high risk of forest loss, the relation between these two variables was evaluated by computing a raster image that measures the distance between the borders of productive areas (agriculture, pasture, coffee, sugar cane, and infrastructure). Forested areas showed a greater probability of deforestation according to their greater proximity to productive areas (Figure 39).

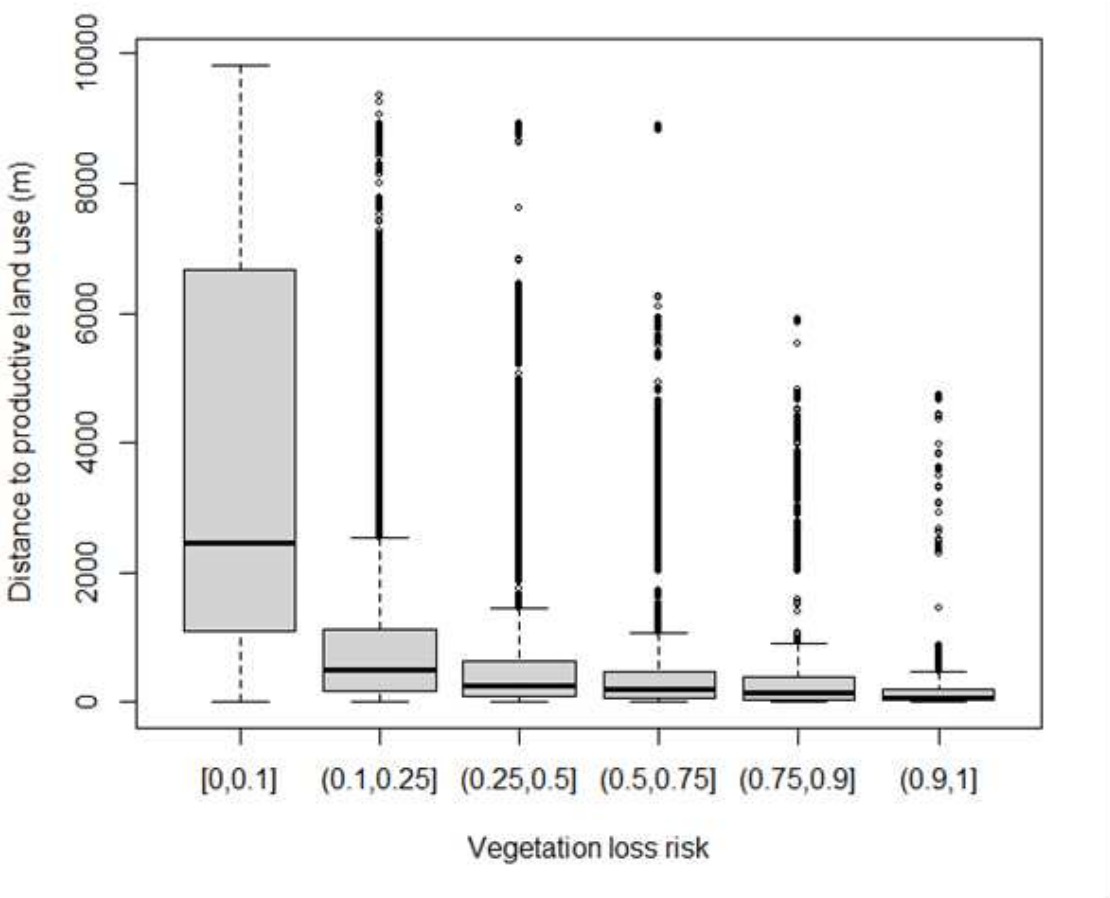


Figure 39. Boxplots showing the decrease in the distances from productive land use with increasing probability risk of forest loss.

4.2.4. Analysis of farms participating or not in the PES program and deforestation risk probability

Farms were classified in those that participated in the PES program (n = 10) at any moment between the start of the program in the year 1997 and the year 2017 and those which did not participate (n = 57). For the comparison, only farms that received payments for forest protection were selected. The number of years the farms were selected to participate in the PES program changed among them, and some of them only participated one year, whereas others participated during several years. The fact that some farms participated more years than others, does not change the interpretation of the results since the predictive model was trained using spatial information on all historical vegetation loss in the area, not taking into account whether the area was included in the PES program or not. Figure 40 shows the distribution of the farms visited, whether they participated in the PES program or not.

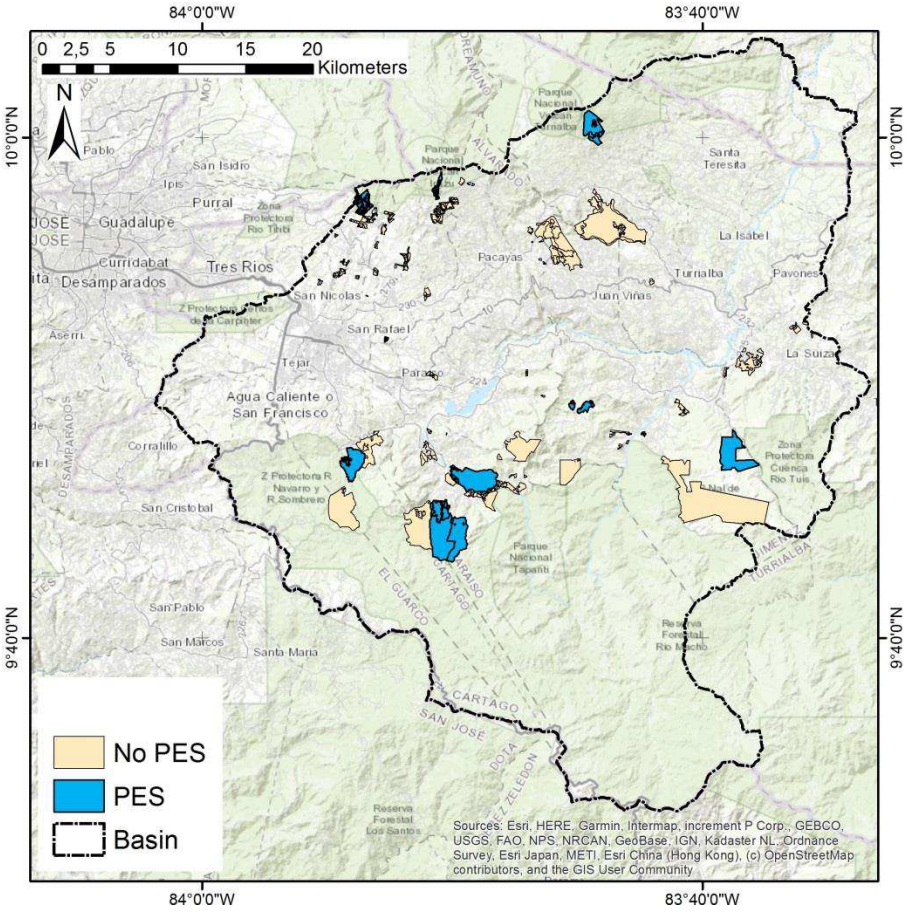


Figure 40. Location of farms participating and not participating in the PES program in the study site.

The risk of forest loss was analyzed for farms participating and not participating in the PES program. For farms participating in the PES program, the minimum forest size for a farm was 15.93 hectares and the risk observed was very low, only 0.27 hectares were found to belong to the risk interval of (0.5-0.75), which represents only 1.46% of the total forested area. The maximum forest size in this group was 482.94 hectares, with also a very low share of forest at risk. Only 4.33% (26.10 hectares) of the forested area showed a loss risk above 0.5 and specifically, only 0.03% (0.18 hectares) of the forested area showed a risk higher than 0.9. In general, for this group forest risk for the highest interval (0.9-1), was very low and below 1% for all of them (Figure 41).

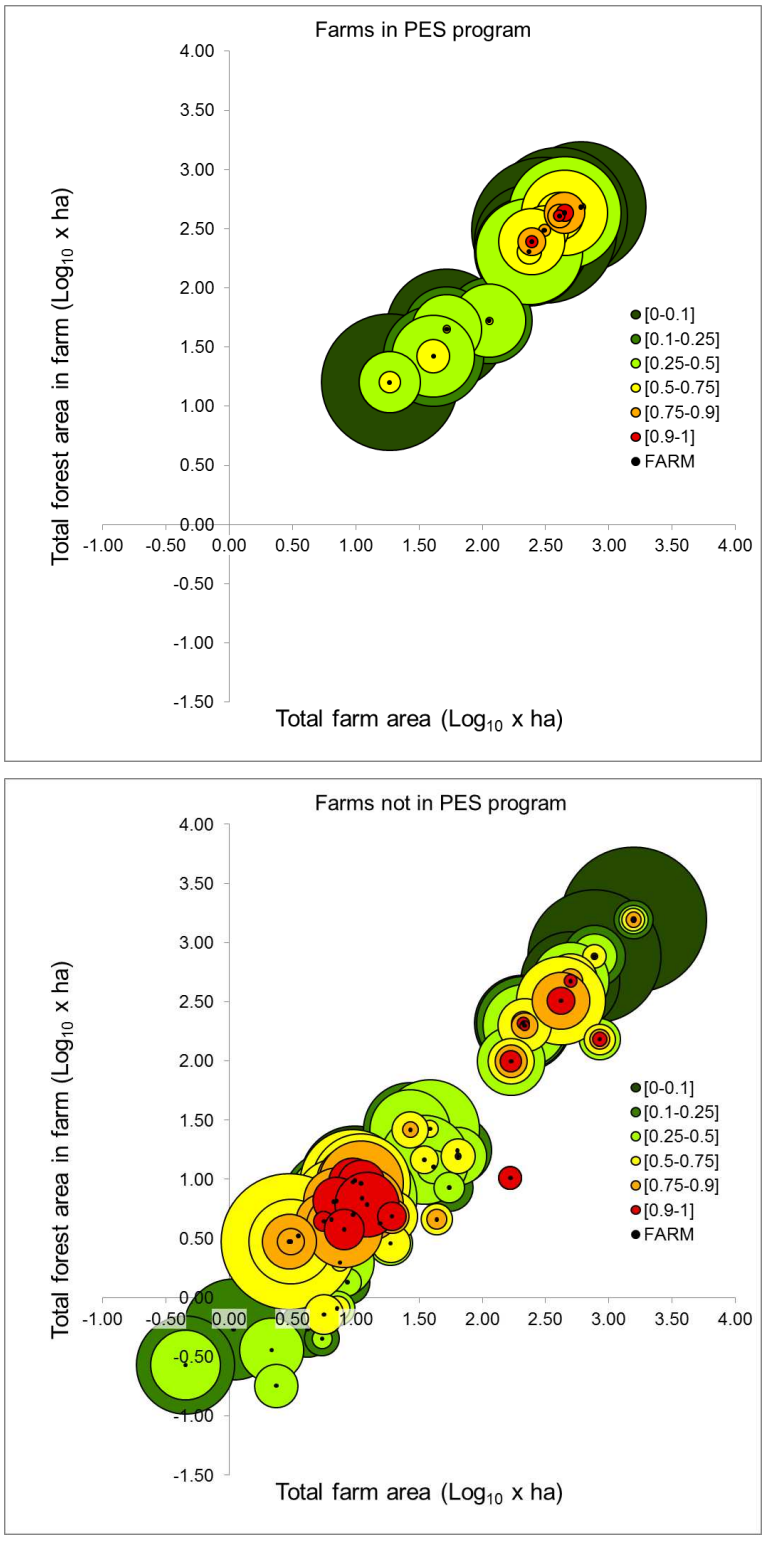


Figure 41. Scatterplot showing the distribution of farms (black dots), based on total farm area and total forest area. Size of the circles correspond to percentage of forest area at each deforestation probability risk interval. Colors correspond to different probability risk intervals. Areas have been transformed using Log_{10} for ease of visualization.

Farms that did not participate in the PES program showed a greater variety of forest sizes within their farms (Figure 41). Forest size ranged from a minimum of 0.18 hectares to a maximum of 1,563.12 hectares. The highest risk of forest loss is observed for a group of farms with forest areas ranging from 3.78 to 9.45 hectares, in farms of different types, but characterized for having productive land uses such as sugar cane or coffee close to the forested areas at risk. Farms with the same characteristics as those receiving PES, this is with forest areas greater than 15.93 hectares, showed similar values in the 0.9-1 risk interval, all below 1% of the forested area, except for two farms, one with 3.07%, which corresponds to a farm with trees waiting to be extracted from a plantation and another with 1.87% of forest area at high risk which could be potentially converted to pasture.

Based on the previous analysis of the linear relationship between the share of forest area within the total area of the farm, a selection of 10 farms that participated in the program and 10 that did not participate was made, to be consistent with the other group. All of them were similar in terms of size of their total forest area and total farm area, and the share of forest area as a percentage of total area. 51% of the total forest area was in the no PES group and 49% in the PES group (Figure 42 and Table 6).

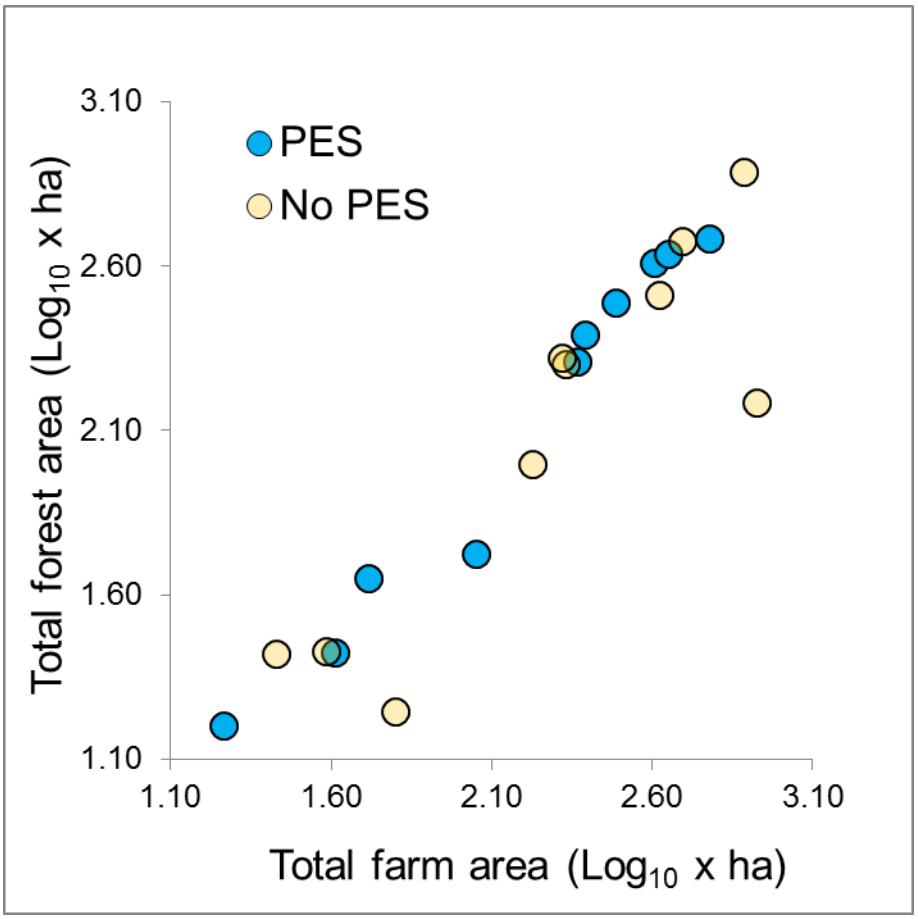


Figure 42. Scatterplot showing the areas of 10 farms participating in the PES program and 10 farms not participating of similar characteristics in terms of total farm area and total forest area. Areas have been transformed using Log_{10} for ease of visualization.

Table 6. Total farm area, total forest area and percentage of forest area of selected similar farms participating and not participating in the PES program. Farms are ordered in decreasing order of total forest area.

No PES			PES		
Total farm area (hectares)	**Total forest area (hectares)**	**Percentage of forest area (%)**	**Total farm area (hectares)**	**Total forest area (hectares)**	**Percentage of forest area (%)**
768.87	766.08	99.64	602.46	482.94	80.16
496.8	473.94	95.40	447.3	430.74	96.30
418.86	323.01	77.12	405.9	405.9	100.00
209.61	208.89	99.66	308.52	307.26	99.59
215.73	199.17	92.32	246.15	245.97	99.93
846.27	152.46	18.02	233.91	202.77	86.69
168.66	99.63	59.07	113.13	52.92	46.78
38.43	26.64	69.32	52.11	44.82	86.01
26.91	26.19	97.32	40.95	26.55	64.84
63.54	17.55	27.62	18.45	15.93	86.34
TOTAL	**2,293.56**			**2,215.80**	

Forest areas belonging to PES and no PES groups were further compared for differences using physical variables such as distance to productive areas, distance to roads, slope, aspect, altitude, or NDVI values (Table 7). The Wilcoxon signed-rank test was used to determine if the differences in their mean values could be considered significant. Statistically significant differences between the median values were found for all variables except for altitude (Figure 43 and Table 8). The results indicate that forests participating in the PES program are located further away from roads and productive land uses and have slightly steeper slopes. The slightly higher difference in altitude could also explain lower values of NDVI for PES forests, which could correspond to mature forests.

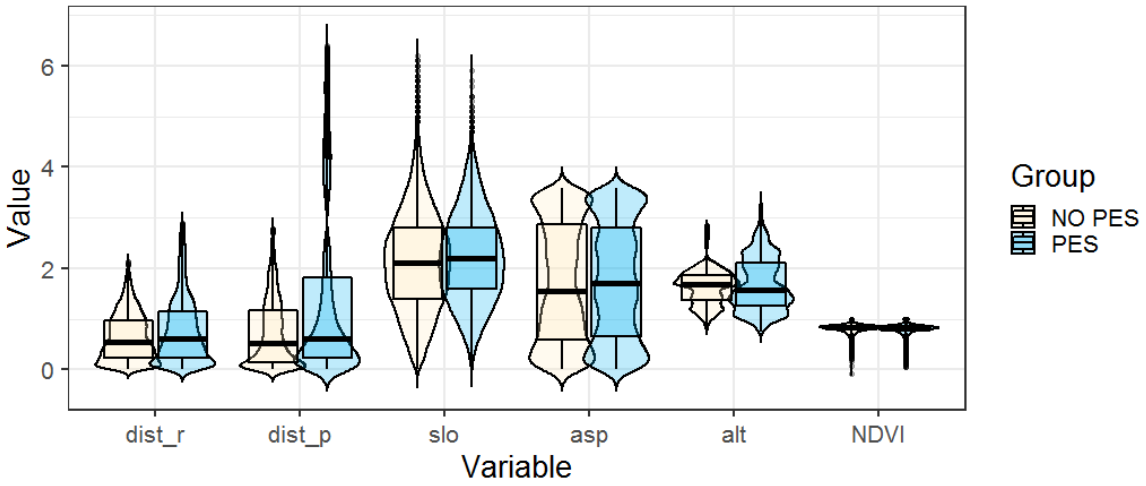


Figure 43. Boxplots and violin plots of the selected physical variables between farms participating and not participating in the PES program. Original units have been transformed for each variable have for ease of visualization (To obtain original units they have to be multiplied except NDVI: distance to roads (dist_r) x 1000 (meters); distance to productive areas (dist_p) x 1000 (meters); slope (slo) x 10 (degrees); aspect (asp) x 10 (degrees); altitude (alt) x 1000 (meters); Normalized Difference Vegetation Index (NDVI) is in original units (adimensional).

Table 7. Summary of forest selected physical variables in farms participating and not participating in the PES program.

	No PES (n=26,048)		**PES (n=24,704)**	
Variable	**Mean**	**SD**	**Mean**	**SD**
Distance to road (Km)	0.63	0.46	0.77	0.64
Diustance to productive land use (Km)	0.71	0.64	1.36	1.62
Slope (º)	21.25	9.69	22.09	8.85
Aspect (º)	170.70	117.88	174.74	114,76
Altitute (meters)	1,632.29	333.67	1,684.84	528.77
NDVI	0.81	0.07	0.80	0.08

Table 8. Results of Wilcoxon signed-rank test for forest variables (p-value < 0.05). Variables

Wilcoxon Signed-Rank Test			
Variable	**W value**	**p-Value**	**Difference between PES and No PES forest areas**
Distance to road (Km)	$2.96x10^{8}$	$2.20x10^{-16}$	Significant
Distance to productive land use (Km)	$2.71x10^{8}$	$2.20x10^{-16}$	
Slope (º)	$3.05x10^{8}$	$2.20x10^{-16}$	
Aspect (º)	$3.15x10^{8}$	$6.08\ x10^{-5}$	
NDVI	$3.41x10^{8}$	$2.20x10^{-16}$	
Altitute (meters)	$3.20x10^{8}$	0.604	Not significant

Differences in the types of land use proportions between the two groups were also analyzed (Figure 44). The areas for all land uses within each group were added and then showed as a proportion of the total area for each group. Farms not participating in the PES program showed increasingly more productive uses, such as coffee, pasture, or agriculture than those participating.

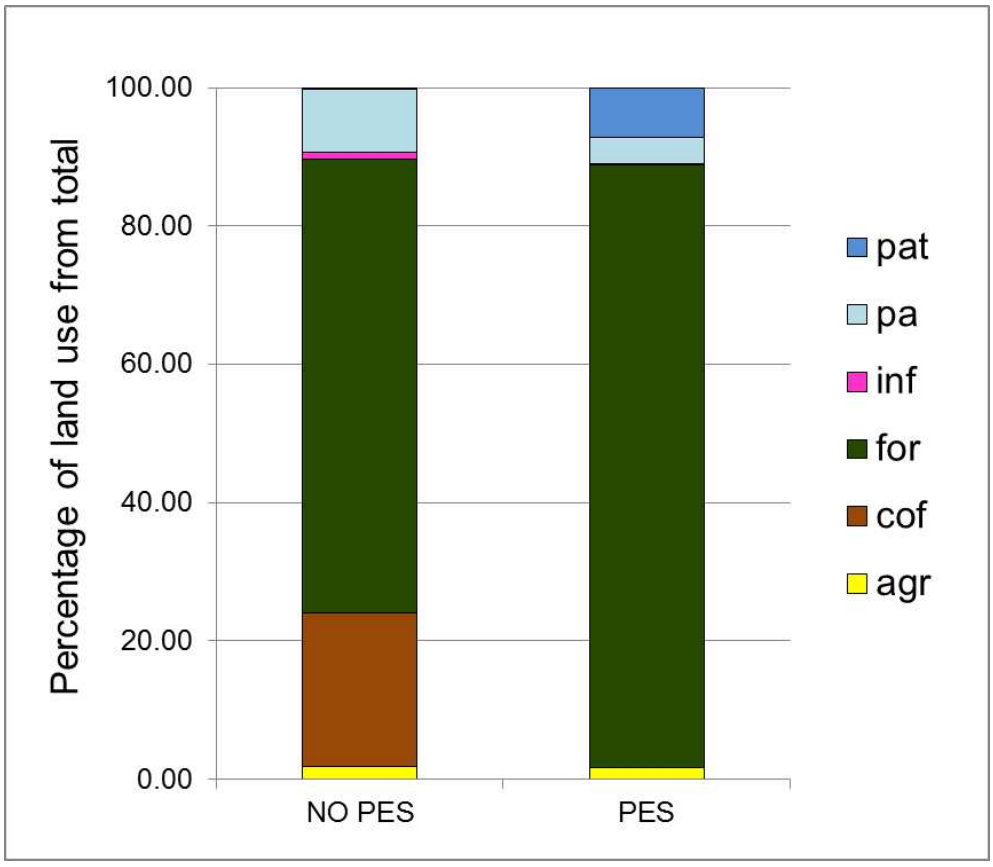


Figure 44. Stacked bar plot showing the sum of all land uses for each group, participating and not participating in the PES program as a percentage of the total area for each group (pat= pasture with trees; pa= pasture; inf= infrastructure; for=forest;cof=coffee and agr=agriculture).

Additionally, a comparison was made between groups of farms participating and not participating in the PES program. The forest area with the lower risk interval (0-0.1) was analyzed in the PES farms, totaling 360.70 hectares. This data was compared with the forest area at the highest risk (interval: 0.9-1) for non-PES farms, with 65.66 hectares. Therefore, in a potential situation, an equivalent of protecting only 18% of the forest area from PES farms at low risk would be enough to compensate the whole area of forest at high risk of deforestation predicted in the non-PES farms (Figure 45).

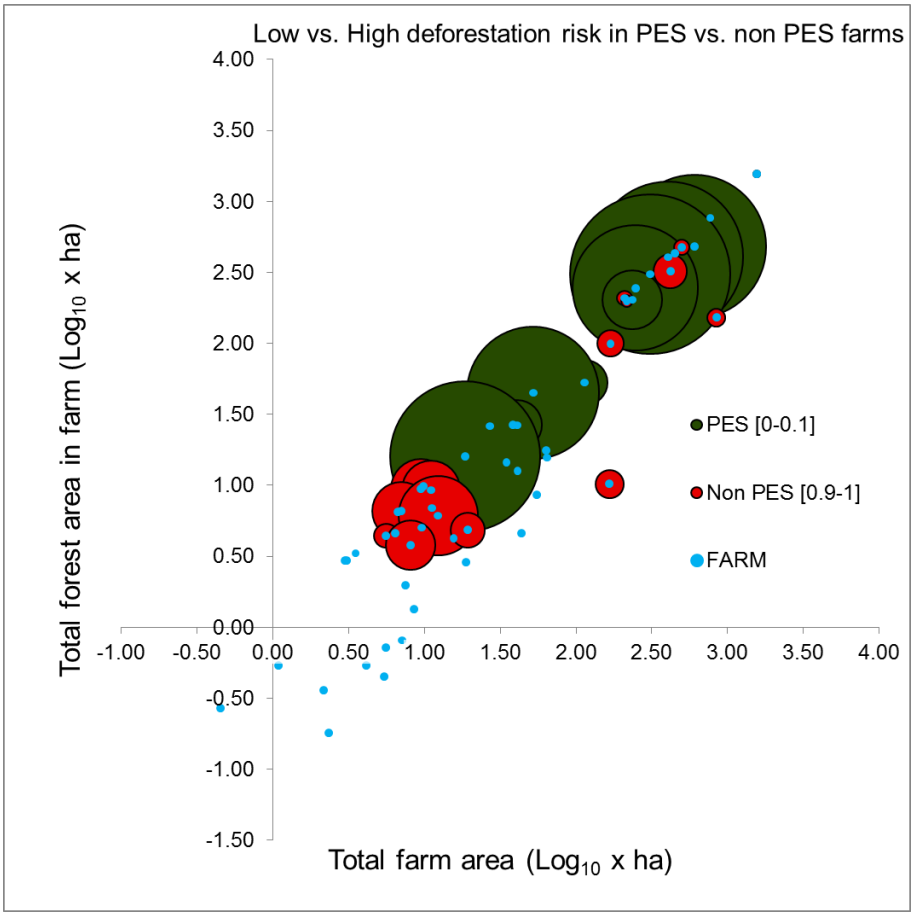


Figure 45. Scatterplot showing the distribution of selected farms (blue dots), based on total farm area and total forest area. The size of the circles corresponds to the percentage of forest area at each deforestation probability risk interval. Colors correspond to different probability risk intervals. Areas have been transformed using Log_{10} for ease of visualization.

The analysis between similar farms in terms of total area and forest area participating and not participating in the PES program also showed low risks of deforestation (Figure 46). From the 10 farms analyzed from each group, the largest areas of forest at deforestation risk were 5.15% and 5.0% for the highest risk interval (0.9-1) for two farms that did not

participate in the PES program. The rest of the farms showed the highest risk (interval 0.9-1) well below these numbers. In third place, a farm not participating in the PES program showed a predicted 3.15% deforestation loss at high risk and all the remaining farms showed values lower than 1% for this interval. 5 farms participating in the PES program showed no forest area at risk for intervals (0.75-0.9) and (0.9-1), only 2 farms in the non-PES group showed similar values. In summary, the majority of the high probability of deforestation risk is predicted to occur in the farms that did not participate in the PES program.

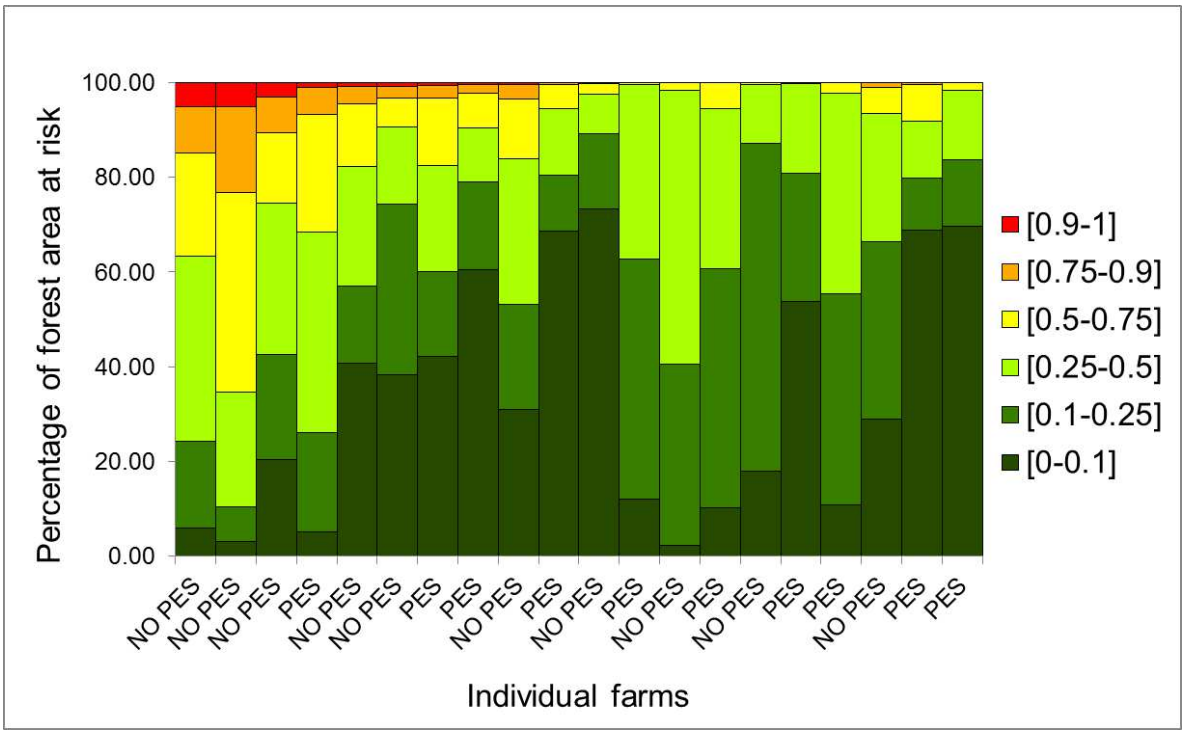


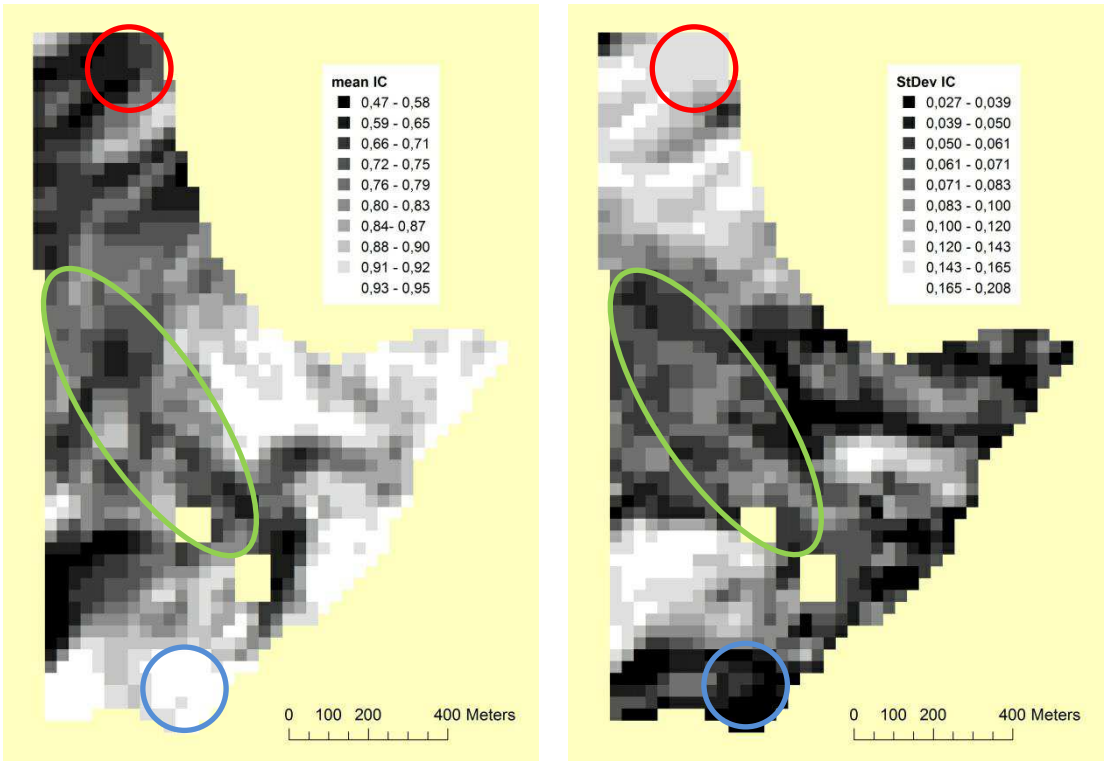
Figure 46. Stacked bar plot showing the risk of deforestation for the selected farms participating and not participating in the PES program

4.3. Analysis of the topography and illumination condition

4.3.1. Illumination condition and vegetation indices

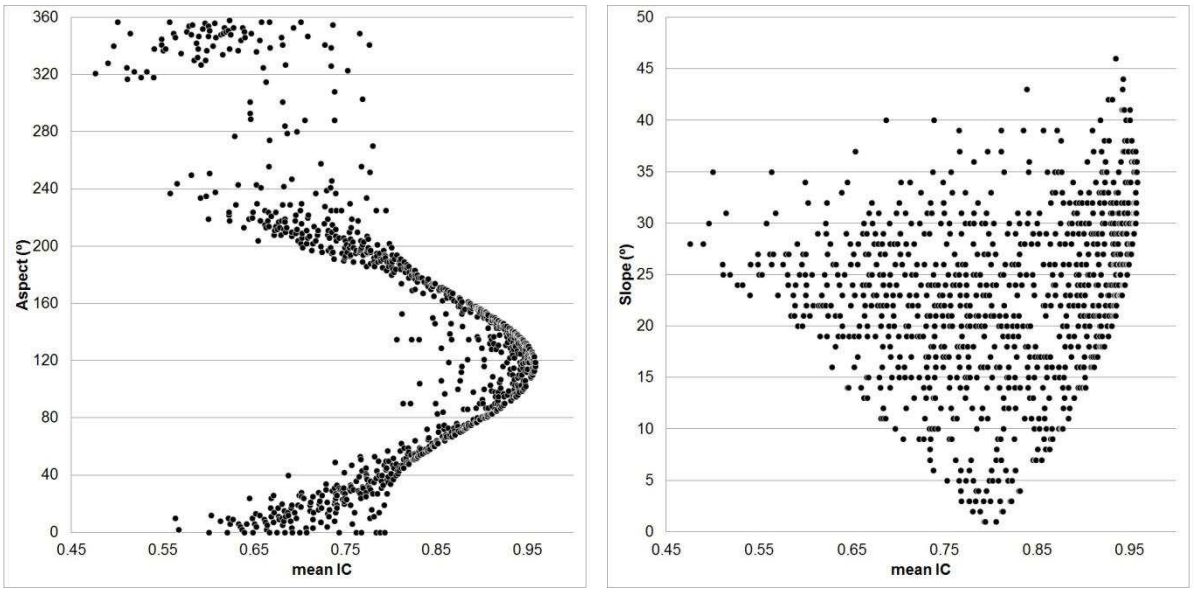
Mean and standard deviation of IC values for the selected images were calculated for each pixel and transformed into an image (Figure 47). Mean IC values ranged from 0.47 to 0.95, whereas the standard deviation ranged from 0.027 to 0.208. Values for both images were unevenly distributed across the image as a result of the effect of topography and sun-sensor geometry. Pixels with low mean IC values correspond to frequently shadowed areas across the 39 images (dark pixels in figure 47a). Due to the sun-sensor geometry, these areas are placed obliquely or even opposite to the sun direction and the radiance received by the sensor is lower or null. Pixels with high mean IC values (bright pixels in figure 47a) are better placed in the sun direction and the radiance received by the sensor is higher. Using the standard deviation as a measure of the magnitude of change in IC for each pixel, it is observed that pixels with high mean IC values experience little variation (dark pixels in figure 47b), compared to those with low mean IC values which experienced greater variation (bright pixels in figure 47b). The visual effect looking at the images is that overall, both mean and standard deviation IC images show opposite patterns, except for an area with low-medium mean IC values and low-medium IC variation (green ellipse in figure 47). This area would correspond to the terrain where the surface is parallel to sunlight beams (Figure 48), therefore, keeping low mean IC values and low variation. Figure 48 shows how mean IC is related to aspect and slope in the terrain. Aspect shows the highest mean IC values in the interval of 80-160º with maximum values around 120º. The lowest mean IC values are in the interval 320-360º and 0-40º, with the minimum around 360º (Figure 48a). Areas with aspect values around 120º would correspond to areas well oriented to direct sunlight, whereas those around 360º would be placed opposite to the sun (shadowed). It is hypothesized that the areas described as parallel to sunlight beams might be those around 200-280º, where low mean and standard deviation IC values are found. These areas can be spatially located in figure 15. The slope shows high mean IC values (~0.8) for flat areas (slope = 0º), and then increasing slope values show a variety of IC depending on how the surface is oriented to the sun (Figure 48b). This effect can be easily seen between the highest mean IC value (0.95) found at 34º of slope and 119º of aspect, whereas the lowest mean IC value (0.47) is found at 28º of slope and 321º of aspect, both slope values are close, but they have opposite orientations (202º of difference). Although variations in the IC have proven to be normal as an effect of

seasonality and sun-sensor geometry variability across the year, it would be expected to find IC variations with a similar degree of change across the whole image. These results suggest that variations depend on terrain conditions (Figure 48). However, if it is considered that terrain, topography, and cover are constant, the uneven variation in different IC values must be affected by changes in sun-sensor geometry. To investigate the magnitude of these changes and their influence on VI, the correlations between VI~IC were computed.



(**a**) (**b**)

Figure 47. Spatial distribution of the mean illumination condition (a) and its standard deviation in the area (b). Each pixel shows the result of the 39 selected images. Red and blue circles show areas with opposite patterns. The green ellipses show areas with low mean IC and variation.



(**a**) (**b**)

Figure 48. Scatter plot showing mean IC and aspect (º) in (a) and IC and slope (º) in (b) at the pixel level.

The analysis of the Pearson correlation between IC and VI on a pixel-basis showed both negative and positive trends in the area clearly associated with IC. The correlation EVI~IC showed a higher range of values (-0.55-0.97) than the correlation NDVI~IC (-0.25-0.73), confirming that EVI is more sensitive than NDVI to changes in IC. Positive correlation values were more abundant than negative ones, describing overall positive trends for both VI, meaning that for an observed increase in IC, VI also increases. However, not all correlations were significant. The EVI~IC correlation was significant ($p<0.05$) with positive correlations for most of the pixels and negative correlations for a few pixels (Figures 49 and 50). Areas with positive EVI~IC correlations correspond to pixels with low mean IC in figure 47a, suggesting that EVI is very sensitive to increase when IC increases in poorly illuminated areas, showing the strongest correlation in pixels with the lowest IC (Figures 48a, 48b). A small number of pixels showed significant values for negative correlations between EVI~IC in well-illuminated areas. This can be related to the appearance of tree shadows between canopy levels when IC conditions change, which has been found in other studies (Brede et al., 2015). The response of EVI to trees shadowed by other trees would be lower. However, higher resolution imagery would be needed to describe the canopy structure and the distribution of canopy shadows with different IC to account for this effect. The correlation NDVI~IC was significant only for some areas that showed positive correlations and not in areas with negative or no correlations (Figures 49 and 50). The strength of the positive correlations NDVI~IC was also lower compared to those in

EVI~IC. The increase of NDVI with increased IC was also found in pixels with low mean IC and low IC variation, previously described as having a parallel orientation to that of the sunlight and limited to a well-defined aspect range. It is known that NDVI saturates rapidly with high biomass levels and dense canopies and is less sensitive to changes in IC, which would explain the low correlations NDVI~IC found (Matsushita et al., 2007; Peng et al., 2018). However, the fact that NDVI seems to be sensitive to changes in areas with low mean IC and low IC variation is a new finding that needs to be further investigated. Because the correlation is calculated with all IC and VI values available for each pixel and these pixels correspond to different dates in different seasons and years, the correlations for both VI are reflecting intrinsic temporal patterns associated with IC.

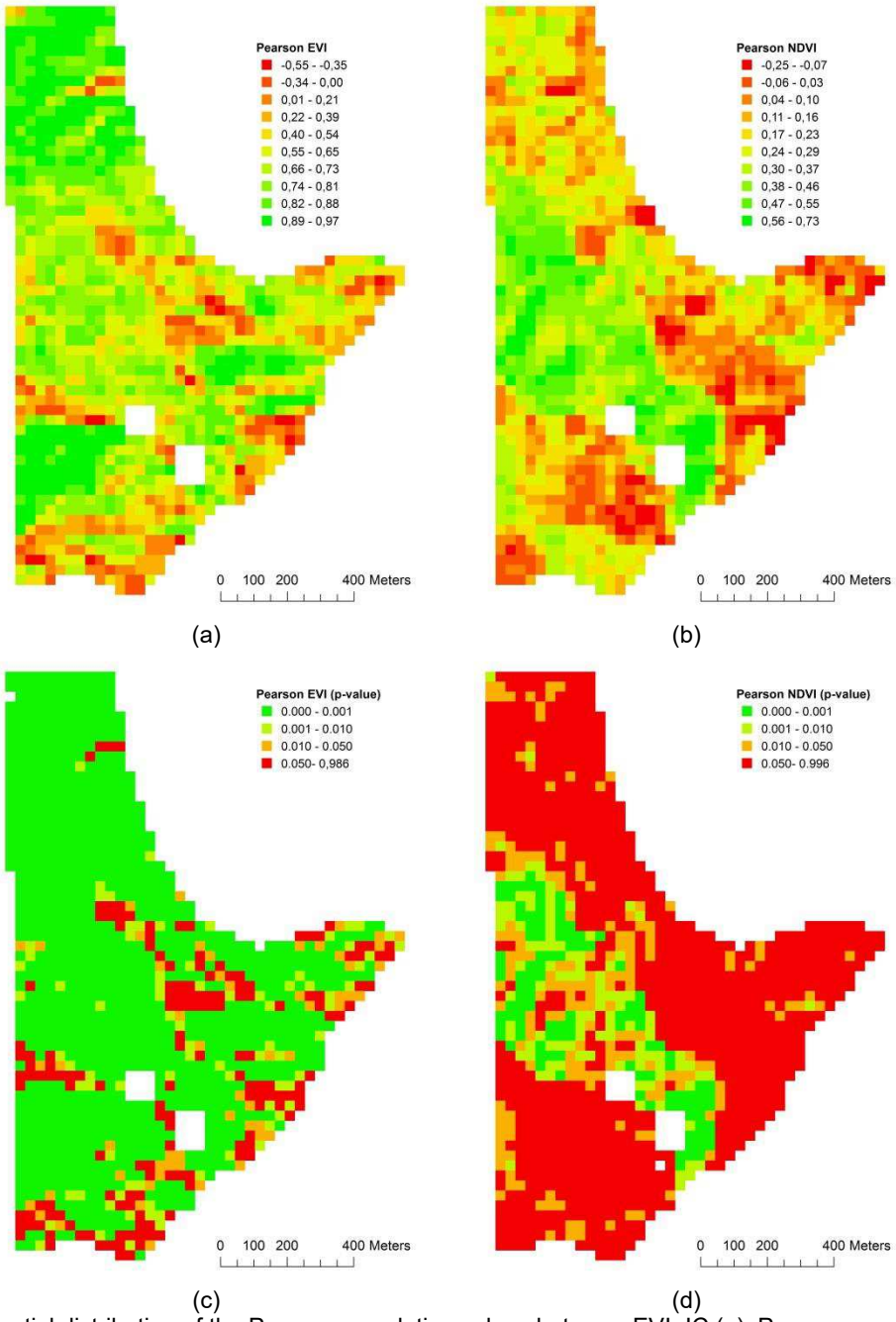
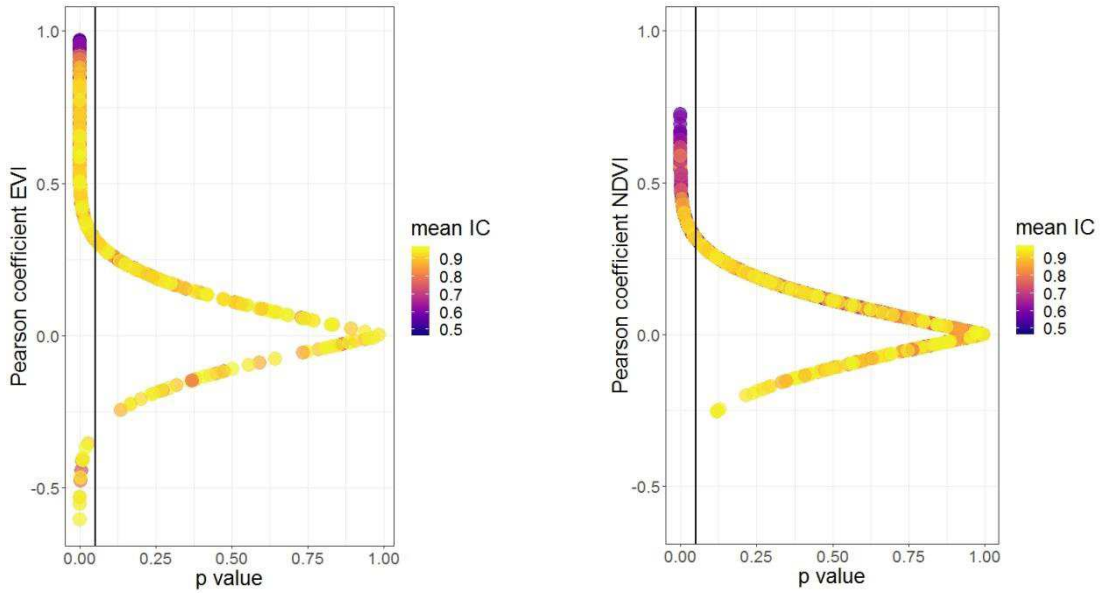


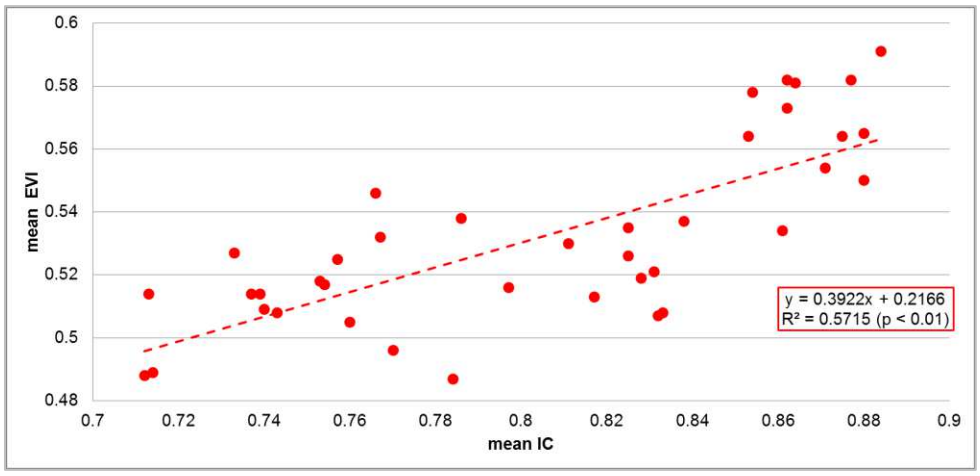
Figure 49. Spatial distribution of the Pearson correlation values between EVI~IC (a); Pearson correlation values between NDVI~IC (b); Significance values ($p<0.05$) for Pearson correlation between EVI~IC (c); and Significance values ($p<0.05$) for Pearson correlation between NDVI~IC (d). Each pixel shows the result of the 39 selected images.



(**a**) (**b**)

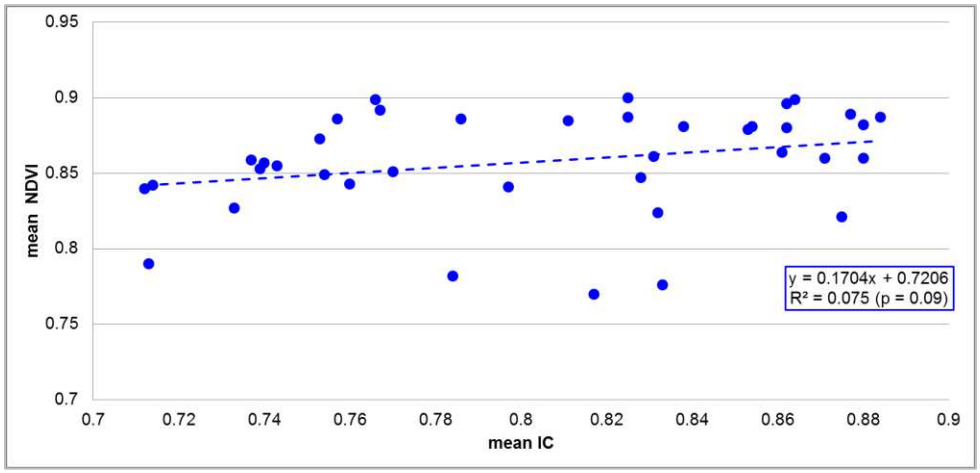
Figure 50. Scatter plot showing p values and Pearson correlation coefficient values for the correlation EVI~IC (a) and NDVI~IC (b). The black vertical line represents p = 0.05, being all points to the left significant at p<0.05. Colors represent low mean IC values (purple) to high mean IC values (yellow).

When the relationship between EVI~IC and NDVI~IC is described comparing the mean values for each selected image (n=39), EVI and IC showed a significant and positive dependence (R^2=0.5715, p <0.01) (Figure 51), whereas NDVI and IC showed a very small to almost no dependence (R^2=0.075, p <0.09) (Figure 52)**.** While these results corroborate the higher sensitivity of EVI and the lower sensitivity of NDVI to the effect of IC, they are not able to explain alone the VI~IC variations found when we examine the image at the pixel level.



(**a**)

Figure 51. Scatter plot of the mean illumination condition (IC) versus the mean EVI value for each of the 39 selected images .



(**b**)

Figure 52. Scatter plot of the mean illumination condition (IC) versus the mean NDVI value for each of the 39 selected images.

Furthermore, the images used in this analysis correspond to different dates in different years. Because it is known from other studies that IC change at different dates (Galvão et al., 2011; Maeda et al., 2014; Maeda and Galvão, 2015; Morton et al., 2014; Ponzoni et al., 2010), the images were chronologically ordered to investigate temporal trends on IC in the next section 3.2.

4.3.2. Time series for IC, EVI, and NDVI from 1984-2017

Temporal trends for IC, EVI, and NDVI show positive and similar slopes and are highly significant ($p<0.01$) (Figure 53). NDVI shows the best adjustment ($R^2=0.57$, $p<0.01$) followed by EVI ($R^2=0.54$, $p<0.01$) and then by IC ($R^2=0.28$, $p<0.01$), the latter which showed the poorest adjustment due to the high variation as a result of known seasonal changes. For the 39 selected images, the minimum mean IC value (0.71) corresponds to the date 31/12/1989 with mean EVI and NDVI values of 0.48 and 0.84 respectively. The maximum mean IC value (0.88), corresponds to the date 04/09/2016 with mean EVI and NDVI values of 0.59 and 0.88 respectively. A comparison between two close images, both from the dry season but separated 28 years also depicts the effect of the increase of mean IC in both mean EVI and NDVI (Figure 54).

Observing an increase in IC, EVI, and NDVI over time despite having a limited number of images but still representative of different seasons and years to make this analysis would

also indicate that there is a general improvement in IC and hence in EVI and NDVI from old to new Landsat images.

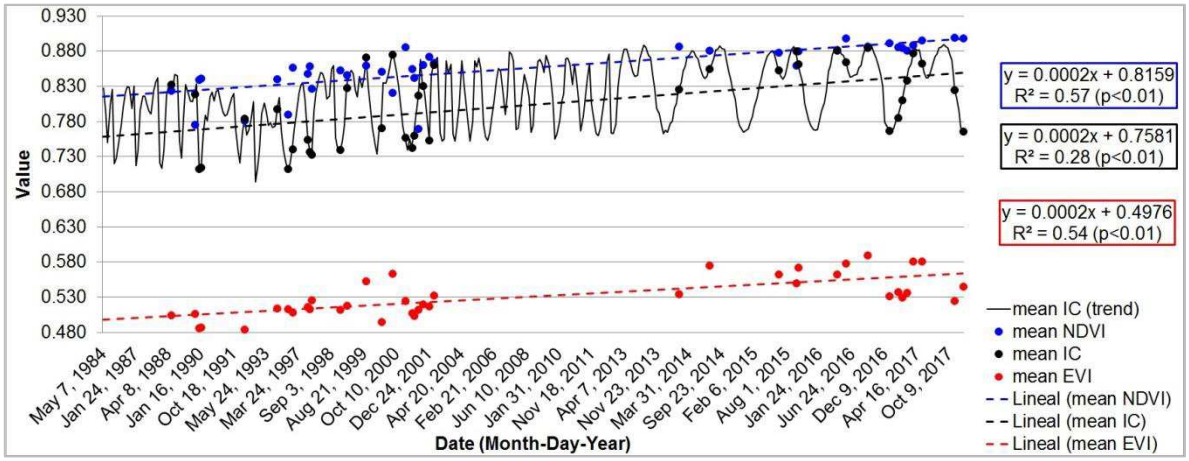


Figure 53. Time-series plot of the mean IC, mean EVI, and mean NDVI values for all the 39 selected images. The solid black line represents the mean IC calculated for all the images available (n=397).

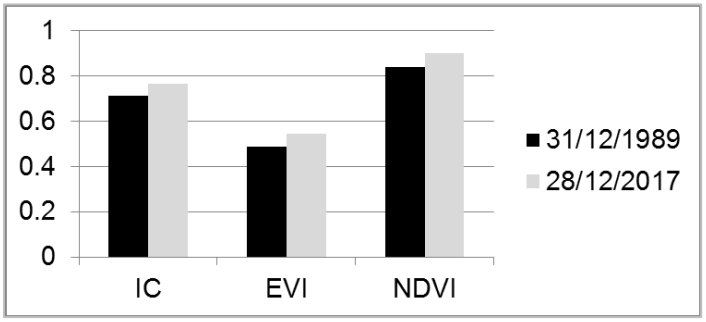


Figure 54. Barplot comparing mean values of IC, EVI, and NDVI for two close images in date but 28 years apart.

The modeling of IC for all images in the period studied is shown in figure 53. The mean IC value in the area shows an overall increase over time. This can be observed comparing images close in date but separated in years, where the increase in IC occurs to a similar degree for every date and month compared (Figure 55). In more detail, the mean IC trend experiences an overall increase with a slight drop between the years 1989-1993 and then an increase again. It can be also observed that due to missing images in some dates between the years 1984-2012, the IC seasonality pattern cannot be as clearly seen as after the year 2012 to present when the coverage of images is more complete. Overall, EVI and NDVI show higher values in the recent period 2014-2017 than in the previous

period 1988-2002 regardless of the seasonal effects. The lack of clear images between the period 2003-2013 is due partly to cloudiness and to the Scan Line Corrector (SLC) failure in Landsat 7 (USGS, 2019c).

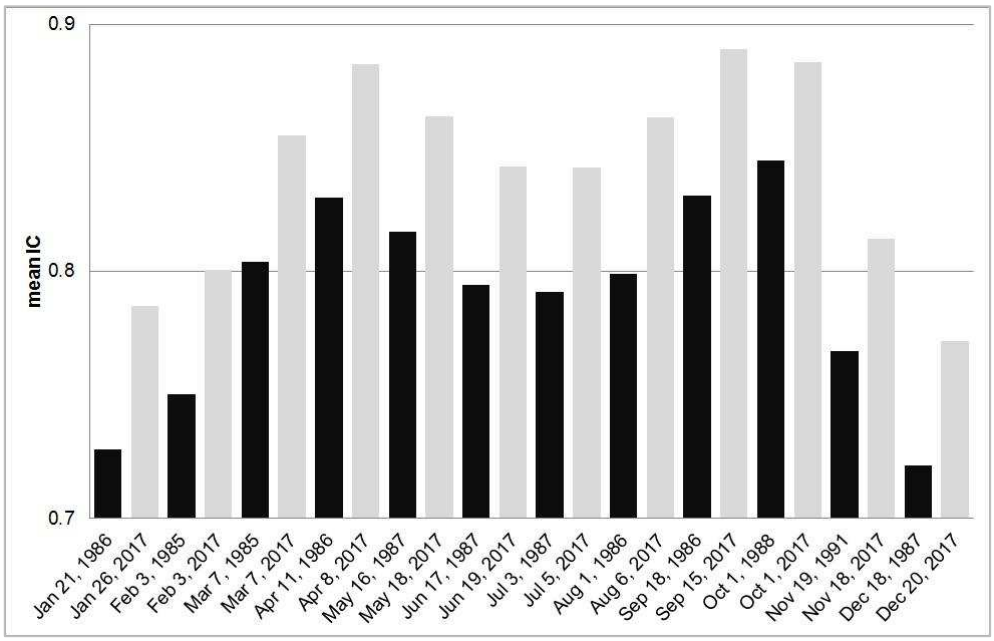


Figure 55. Comparison of mean IC for images close in date but separated in years. The criteria of selection were first close dates, then the highest number of years in difference.

Monthly trends are shown in Figure 56. Clear seasonal trends can be observed. The lowest mean IC occurs in the dry season (December-February) and higher IC can take place at two different times in the year, at the beginning of the rainy season (April-May) and in the short dry period within the rainy season (August-September). Every month shows a similar and steady increase in IC over time from old to recently acquired images, although the strength of the increase varies between months. March shows the best adjustment in the increase of IC with time (R^2=0.71), whereas November showed the lowest (R^2=0.18). Besides the drops in IC between 1989-1993 that were commented on previously and that can also be seen in the figure, anomalous changes are occurring in December for the latest imagery available. However, this behavior would require further study. The magnitude of the increase reflects constant trends across years for all months. Although beyond the scope of this study, further research would be needed to explore these trends with phenology patterns.

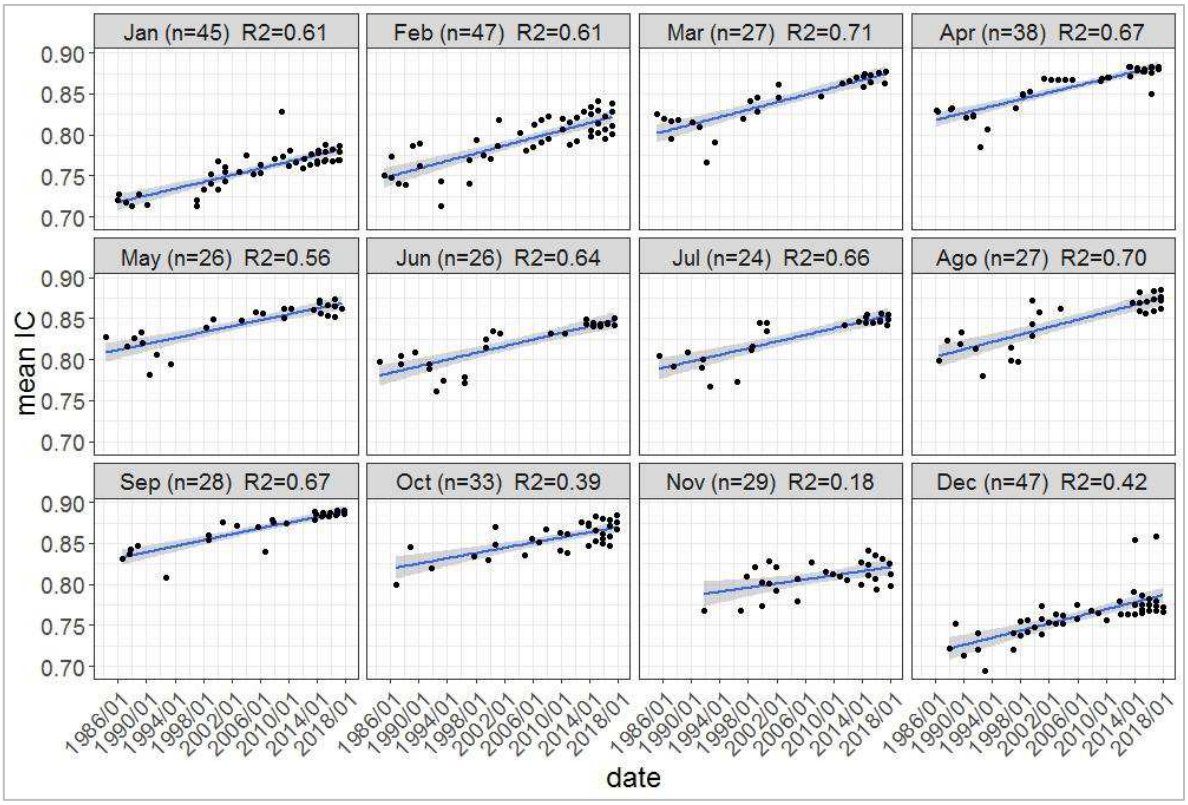


Figure 56. Temporal trend of mean IC in the study area by month and year for all Landsat images selected (n=397). Blue lines represent regression lines and grey areas CI at 95%. The number of images available for each month(n) and R2 are displayed.

Figure 57 shows the mean IC for all Landsat images in the area in chronological order indicating the sensor capturing the scene. Most of Landsat 4 and 5 images covered the period 1987-2001 and then Landsat 5 alone covered some dates between 2007-2010. The period for Landsat 7 ranges from 1999-2017 with a period of synchronicity with Landsat 8 between 2013-2017 that has produced more frequently available images during this time. The lowest mean IC value (0.69) was found on 31/12/1992 using Landsat 4 and the highest (0.89) on 15/09/2017 using Landsat 7. More detailed analysis also shows improved IC within the same sensor. For a variety of close dates in different years, Landsat 5 and Landsat 7 show a general increase in IC.

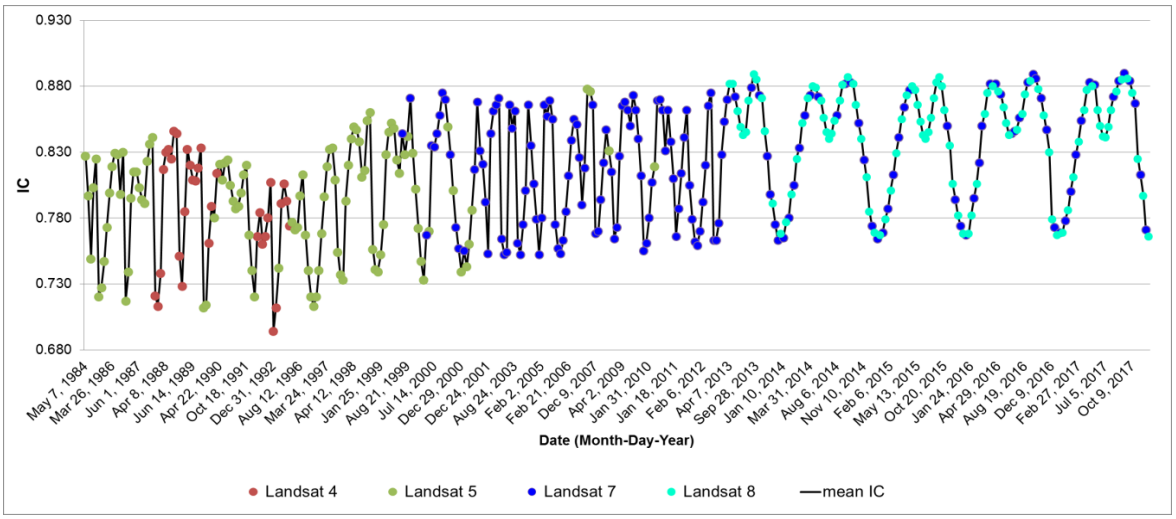


Figure 57. Temporal trend of IC in the study area for all Landsat images selected (n=397) and all Landsat sensors.

4.4. Farmers'perception

4.4.1. Farmers' perception of the forest and the environment

In general, respondents had a very positive perception of the forest in the area and overwhelmingly associated the presence of trees with the presence of water.

"Without the forest, there is no water", "without forest, there is no rain" [Respondents 2 and 3].

Owners that had water springs inside their farms explained the importance of protecting them with trees or keeping the forest around them. Respondents also associated the logging of forests in past times with the general loss of water and the dry-up of rivers that they experience today.

The next most repeated benefit of forests was the production of oxygen, and their role as air purifiers so that trees were broadly perceived as "lungs". Additional features of forests were their role as temperature regulators,

"If we stay in the middle of the square, it is so sunny that we are very hot, whereas in the forest with many trees is way different" [Respondent 4].

Concerning this effect, trees were also good for shadowing the cattle in open pastures or pastures with scattered trees. Usually, these pastures are located at high elevations and sun radiation can harm the animals, some respondents said. Grass growing under the trees is also highly appreciated by the cattle. Additional benefits of forest trees were their role as windbreakers, especially in the boundaries of agricultural areas; their use for firewood/construction, but only for dead trees; their role in attracting birds and biodiversity, and the well-being and health that their presence provoke.

However, perceptions changed and the benefits of trees were prioritized between the different landscapes depending on the most predominant land uses across the basin. In the agricultural area, under the Irazú volcano, where trees are scarce and farmers are suffering droughts in recent years, the perception was that more trees were needed in the area. This would help to increase the quantity of water needed for irrigation in the area. A former hydrologist from the Costa Rican Institute of Electricity (ICE) and farmer, suggested that “the PES program should promote tree conservation in agricultural areas”.
Among tree species, cypress trees were not considered good and farmers considered it correct to cut them in favor of native species. Plantation of cypress trees was the only alternative found to stabilize the soils surrounding Irazú volcano after the ash eruption of 1963 to prevent ash and mud floodings in the city of Cartago. After this period, the cypress plantations are old and some trees are dying and falling, and the Government has started actions to replace them with native species. Farmers in the surrounds of Irazú volcano complained about the abundance and low utility of cypress trees in the area, the wood was not good nor the branches as fodder for cattle consumption.
In the area of sugar cane, where the forest is predominant in the surroundings and the landscape is densely covered by trees, farmers also acknowledged the importance and benefits of trees but sometimes commented that some areas could be logged to extend their plantations.

During the interviews, in general, forest loss was not perceived as a risk or a pressing environmental problem. The observed trees in the landscape were maintained for the mentioned reasons and environmental benefits. Only one farm was found to log planted eucalyptus and cypress species, close to Tapantí National Park as it implies an important income for the economy of that farm. The owner asked for the advice of FONAFIFO staff and forest managers to manage and decide over those plantations. Farmers frequently

responded that the 1996 forest law penalizes tree logging and might be a principal reason to prevent not only tree cutting but also poaching, or even profit from dead trees:

"A tree fell and the park rangers did not let them make boards for a bridge, it is illogical" [Respondent 5].

Pressing environmental problems perceived by farmers in order of importance were: 1) extended drought periods and irrigation problems, 2) soil erosion due to cultivation in steep slopes, 3) pollution of water streams due to overuse of agricultural chemicals, and 4) waste pollution in water and the environment as a result of the increase of tourism in recent years. The fieldwork revealed important environmental problems related to agricultural overexploitation of the north and especially the northwestern part of the basin under the Irazú volcano. It was witnessed intensive cultivation of crops (potatoes, onions, or carrots) at steep slopes and high altitudes taking advantage of the rich and deep ash mineral soil in the area. The combination of steep slopes and scarce tree coverage make this area prone to high soil erosion during the rainy season. Every year, the Costarican Institute of Electricity (ICE) faces serious problems due to the amount of eroded soil that ends up in the Cachí reservoir at the bottom of the basin. The reservoir has to be emptied of mud and generate electricity using oil, incurring high economic and environmental costs [Respondent 5]. ICE has continuously provided farmers with trees trying to increase the tree coverage in agricultural areas. Around 2014 ICE had provided farmers with 18.3 million trees to recover and protect the areas around their hydroelectric projects since the first tree was given, in 1960 (Martínez, 2014).
In the southern part of the basin close to Orosí and the Tapantí National Park, soil erosion is attenuated due to the high tree coverage. In this area, the environmental problems are related to garbage and river pollution due mainly to the increasing tourist pressure in the area.

4.4.2. Farmers' perception of the PES program

Although the PES program is known for most of the farmers interviewed, the majority of them expressed a lack of clear information about the program in terms of administrative steps, amounts paid, the minimum size of the area, or other requisites. Some of the farmers that decided to participate also did not found clear rules or information on how to participate in the program. Some of them were dubious about the technical capacity of

government staff in charge of the PES program. Access to information is considered weak for the PES program becoming a factor that limits the interest to participate.

“They are not clear about what they are selling” [Respondent 6]

“Excessive bureaucracy, I am not interested” [Respondent 7]

Farmers perceived a lack of commitment between them and the administration staff. Potential participants expect a personal treatment, each farm has different features in terms of size, land use share or other features such as legal status and this affects the level of workload assumed by the administration. It seemed to the respondents that more complex farms in terms of workload were left aside and this affected their confidence towards the program.

The participation of third parties that are needed to be included in the process is another factor of demotivation to participate. A forest manager is needed to visit, design, and digitize the forest area and sometimes a management plan. This is costly and it is interesting only if the payment received is enough to cover it because in general owners try to use the received payments to cover all the costs. Based on the interviews the manager can charge around 10% of the payment to the owner as an honorarium. Another option is to be an NGO member that can provide these managers and this also comes with an associated periodic fee.

“All this paperwork for so little money is not profitable” [Respondent 8]

The perception of the PES program also changes depending on the type of farmer and the share of forest area inside the farm. Farmers heavily relying on their farms as their main livelihood where most of the area is occupied by productive land uses were distrustful of the PES program and more reluctant to participate. This category of farmers shows a strong conviction that participating in the PES program can affect the use of their lands, which would have a severe economic impact on their activities. Moreover, some farmers located in the agricultural areas are afraid that the government could literally “appropriate” their lands.

On the other hand, farmers located in the southern part of the basin, with a higher proportion of forested areas inside their properties, seem to be more informed about the PES program and with more willingness to participate. These respondents want to include the forested areas in the PES program to be compensated by the environmental services they produce and because the included area in the program can be exonerated from tax payments. The very low opportunity cost of these areas and the restrictions imposed on forest transformation due to the forest law 7575 from 1996, encourage these owners to participate in the program.

“The inclusion of the farm in the program would have helped since it is a non-productive area” [Respondent 9]

Farmers' and forest managers' testimonies gave an idea of the remoteness of some of the farms receiving payments, or how the accessibility would not favor a profitable use of the land nor provide an easy connection to nearby markets or even to urban centers.

“It will take a whole day walking to get to the farm” [Respondent 10].

“The only possible way to get to the farm is by horse” [Respondent 11]

Most of the owners declared that if the payment is obtained, it will be a complement to the activities of the farm, but it will not be enough in any case to substitute the income derived from them.

5. Discussion

5.1. A predictive model for vegetation loss risk: The importance of topography and accessibility

5.1.1. Accuracy and validation of the model

The prediction of the RF model with a 12,80% OOB error rate means that the model is 87.20% accurate in classifying the loss class and therefore in predicting potential vegetation loss. However, the vegetation loss risk map obtained shows a much more extensive predicted area of loss than the historically observed in the years 2000-2018. The reasons why areas at risk predicted by the model were not lost during the historical period may be partly socioeconomic and out of the biophysical aspect of this study but will be discussed in the following sections. These factors, including the effect of protected areas and PES may have prevented these areas to be lost since they have similar features to those found in the historically known areas. The model only predicts suitable areas to be cleared, based on the physical information provided. Another possible reason is that it would be needed more time to reach these areas and clear the vegetation found on them. This reason would explain the OOB error rate in the no loss class of 24.70%, some of the no loss pixels were classified as loss (false positives), because of this reason. The OOB errors observed in this study are, however, in line with other studies that used RF to predict tree species distribution (Evans and Cushman, 2009), mapping of forest, non-forest areas (Mellor et al., 2013) or spatially define areas susceptible to landslides (Shirvani, 2020).

The additional validation using the testing data, with an AUC=0.89, revealed that the predictions are very close to other studies that designed deforestation risk models. Some of them used a RF approach to predict deforestation risk in Borneo (Cushman et al., 2017); in a river basin in India (Saha et al., 2020), or the neighboring country of Nicaragua (Di Lallo et al., 2017). The fact that the high accuracy was obtained using the model on the testing data (n=19,712), which only represents 20% of the initial data available (n=98,564), reinforces the prediction ability and robustness of the model. In summary, RF provided accurate results and proved to have a good predictive power as recently observed in several ecological studies (Di Lallo et al., 2017; Hart et al., 2019; Prasad et al., 2006; Valle et al., 2020).

5.1.2. Importance of predictor variables

Accessibility and topography influence vegetation loss in the study site. Distance to roads is the predominant physical factor associated with deforestation in similar studies across tropical countries (Aguilar-Amuchastegui et al., 2014; Cushman et al., 2017; Di Lallo et al., 2017; Saha et al., 2020; Silva et al., 2020; Zanella et al., 2017). The development of roads was soon related to deforestation in Costa Rica when the transportation routes penetrated the region (Sader and Joyce, 1988). The development of roads is also associated with new settlements and gives better access to heavy machinery, so the appearance of cropland areas or transition from forest to non-forest lands are more likely to occur close to them (Dlamini, 2016; Goldman et al., 2017). The partial dependence plots showed that the probability of vegetation loss is higher within the first 1,000 meters from the main road and then rapidly decreases. road network used in this study was from the year 2010, and the rest of the data were from the year 2000. If the extent of the road network changed enough in 10 years, this might explain in part, the importance of accessibility in predicting vegetation loss. It also highlights that road networks must be as accurate as possible and updated regularly. When designing new targets for forest conservation or protected areas, proximity to roads should be one of the key factors to consider.

Accessibility is also controlled by elevation and slope, being more difficult as these factors increase. The basin is very mountainous and this greatly influences accessibility, for example, there is a maximum difference in altitude of 3,242 meters in only 53 Kilometers across the basin. Elevation also models climate conditions leading to cooler temperatures, has an influence on ecosystems and vegetation distribution, and also influences cloud and rain patterns (Aide et al., 2013; Redo et al., 2012). An example of the conditions created by elevation and slope can be observed in the Tapanti protected area, it occupies the southern part of the basin, its altitude ranges from 700–3,491 meters and precipitation ranges from 2,500 to almost 8,000 mm/year (Bernard et al., 2009). Topographic and environmental variables naturally place physical limits on the types of land use that are possible in an area. The suitability for cropland establishment for example, which is a main driver of deforestation, is highly dependent on environmental and accessibility conditions (Aguilar-Amuchastegui et al., 2014; Bax and Francesconi, 2018; Dlamini, 2016). The decision to clear an area of vegetation to establish a cultivated land is ultimately taken if it conducts to the creation of wealth, therefore, only suitable environments will be chosen

(Redo et al., 2012). In Costa Rica, slope gradient and deforestation keep an inverse relation (Sader and Joyce, 1988), conversely, reforestation has been observed at high elevations (Aide et al., 2013; Redo et al., 2012). Partial dependence plots for elevation and slope showed that vegetation loss presents a U-shaped trend. The probability of vegetation loss is higher at low elevations and gentle slopes. Lowlands are more accessible and appropriate for transport, mechanized agriculture and are warmer in temperature. This area of the basin is occupied by sugar cane plantations, coffee and diverse agricultural products, big cities and settlements are also located in these areas. If land use has to be changed, the economic cost associated will be lower than in more rugged areas. Then for medium values of elevation and slope, vegetation loss is not likely to occur, corresponding with permanently cloudy areas, where the lack of sunlight would not allow the growing of crops. Climate conditions here limit the potential production of land (Bax and Francesconi, 2018). This is the area of the tropical montane cloud forests and its best illustrated with the Tapanti protected area. For very high values of elevation and slope, the probability of vegetation loss increases again, but two factors can be responsible for this. First, some of the vegetation loss observed at high elevation (3,200 meters) is due to the Turrialba volcano ash eruption in 2014, with consecutive eruptions throughout until the year 2016. The ash has covered an extensive portion of vegetation around the crater and it has been classified as loss in the historical loss maps. A second factor is frequently observed landslides produced at high elevations and steep slopes, after heavy rain events and storms, high elevation areas trigger weathering and erosion (Bera et al., 2020).

The evaluation of the indices showed that LSWI had an important effect on the accuracy of the model. LSWI measures liquid water in vegetation canopies and is also sensitive to the total amount of liquid water in the crops. In comparison with NDVI, is better at detecting moisture and does not saturate with high values, so usually, LSWI shows lower values than NDVI for the same type of vegetation (Chandrasekar et al., 2010; Gao, 1996). Negative values of LSWI correspond to bare or very dry soils and values from 0-0.3 correspond to very moist areas with high vegetation cover, at very high values around 0.4, LSWI represents water-saturated vegetation, or flooded areas, like shores of lakes. Therefore, the risk of vegetation loss would be low in areas with high vegetation cover, like dense cloud forests, but would increase as we approach less densely vegetated areas, forest edges, or dry soils, so close to cropland, cropland with scattered trees or open

woodlands and also around areas with water, mainly lakes and reservoirs, where the loss of vegetation can be attributed to flooding. Other studies indicated that deforestation probability increases close to forest edges and decreases far from these areas (Saha et al., 2020).

The findings concerning accessibility and topography, highlight the importance of the application of this method at the local level. Due to the mountainous character of the study area, it has been demonstrated that this factor is key to explain deforestation in this location, but it does not mean that it would be the case in other parts of the country. This would be in line with other studies that found that the variability of topography across the country, impact micro-meteorological dynamics, soil, and human activities, therefore contributing to changing deforestation intensities as these factors change due to topography (Stan and Sanchez-Azofeifa, 2019).

5.1.3. Analysis of historical vegetation loss and predicted vegetation/forest risk loss in PWA and farms

5.1.3.1. Historical vegetation loss in PWA

The historical vegetation loss rate observed in the area can be considered low (-0.14% y^{-1}). During the period 2000-2015 countries with the highest deforestation rates reported -0.35% y^{-1} (Brazil) or -0.56% y^{-1} (Indonesia), whereas Costa Rica showed a net increase in its forest area of 1.06% y^{-1} (Keenan et al., 2015). Other studies indicate that this particular basin must be experiencing a forest recovery from 1986-2008, after years of experiencing forest loss (Stan and Sanchez-Azofeifa, 2019; Vallet et al., 2016).

The rate is even lower if it is considered that this analysis includes additional vegetation other than forest. In terms of size and magnitude, most of the observed loss occurred outside protected areas. Field observations confirmed that while part of this loss affected forested areas in the year 2000, some other areas corresponded to coffee removals due to rotations, sugar cane harvesting, or even flooded areas due to dam construction or reservoir filling (Bonilla-Bonitilla Protected Wildlife Area). At high altitudes, the Turrialba volcano also contributed to the loss of vegetation due to ash release. Vegetation loss was very low for all PWA (<5%) with the only exception of Turrialba volcano.

5.1.3.2. Predicted vegetation loss in PWA

The predicted vegetation loss in the area greatly exceeds the area of historical observed loss. In the study area, 2.83% of the initial vegetation cover in the year 2000 has been historically lost. The model predicts that excluding this already observed loss, about 28.51% of the initial vegetation in the year 2000 will change to other land use. One main reason to explain this difference could be that not enough time has passed for all this change to occur. The model is predicting the locations of potential vegetation loss, but not their loss rate. Using the loss rates observed (-0.14% y^{-1}) and considering that observed past trends will be the same in the future (Goldman et al., 2017), it will take 196 years for this change to happen. Supported by comparisons of general deforestation rates in other tropical countries (Keenan et al., 2015), low changes observed in the same basin by other authors (Vallet et al., 2016), and future projections that predict a net forest recovery for Costa Rica (Stan and Sanchez-Azofeifa, 2019), it is assumed that the rates are low. The worst scenario of future deforestation predicted by Stan and Sanchez-Azofeifa (2019) for Costa Rica during 2013-2068 (-0.13% y^{-1}) is very close to the observed 2000-2018 historical loss of -0.14% y^{-1} in the area. Additionally, and considering the probability values obtained by the model, the areas at risk are even smaller, with only 5,97% of the vegetation at high risk, and when looking at these areas spatially, most of them are located around or very close to already lost patches. The highest uncertainty and misclassification in the model (probability around 0.5) is observed in the transition areas between vegetation and no vegetation, which spatially correspond to areas that are far enough from roads but not as far as remote areas. Similar uncertainty values have been observed when modeling deforestation in Nicaragua (Di Lallo et al., 2017).

One important result of the model is that the most affected areas are located outside the PWA. This would be a natural reflection and would indicate that the designated protected areas are being efficient in their purpose, but this also questions whether protection would be needed in other areas. Following the concept of additionality (Wunder, 2007), is there a difference in terms of vegetation loss between the presence or the absence of the PWA?. The model provides spatially explicit evidence that in general, and based on historical observations, the risk of predicted vegetation loss for all PWA is low. These results confirm previous studies on the effect of PWA location on preventing deforestation during the years 1986-1997 in Costa Rica (Pfaff et al., 2009b). A difference in avoided deforestation

was observed in PWA close to cities, national roads, and lowlands, but almost no deforestation was observed in PWA located far away from urban areas or in rugged areas. This would confirm that PWA in Costa Rica is placed towards areas that face lower deforestation threats (Pfaff et al., 2009b). Steeps slopes, poor accessibility, and low agricultural suitability are factors strongly associated with protection, not only in Costa Rica but worldwide (Joppa and Pfaff, 2010).

The PWA that shows the highest loss risk in percentage is Bonilla-Bonitilla, but the reason for this is simple, it is a wetland, and observed and predicted vegetation loss indicates flooding of the shores and consequent loss of vegetation. The following ones are Guayabo and F. Estado which are next to each other and crossed by a road, and for this reason, the model predicts potential loss. However, Guayabo is an archeological site, and it is unlikely that vegetation loss would be extensive in the future unless the site is expanded due to new excavations. Carpintera is located between Cartago and San José, and therefore, the pressure is high, with potential areas to be converted into pastures if needed. C.V. Central and Navarro-Sombrero are partially occupied by farms and agroforestry systems, so the leap to possible land-use changes may be feasible. Turrialba volcano will continue losing vegetation if ash eruptions continue, but the restrictions of land-use change being a national park are restricted otherwise (Pfaff et al., 2009a).

Additionally, while PWA seems to face a negligible risk of deforestation, especially the most restrictive ones such as national parks and biological reserves, proximate areas have shown different trends, and this would indicate that the area of PWA should be eventually extended (Sanchez-Azofeifa et al., 2003).

Between 1960 and 1997, only 7-9% of protected forests in Costa Rica would have been deforested in the absence of protection (Andam et al., 2008). The model predicts that only 4.43% of the initial vegetation in the year 2000 will be lost for all PWA (this includes observed lost areas), which halves the previous estimates in this period from the year 2000 onwards.

Areas outside the PWA with predicted high risk, are from west to east, patches of trees in urban environments, and lands transformed to greenhouses between the cities of Cartago and Paraíso. At the center, around the Cachí reservoir, most of the lost and predicted areas correspond to either flooded sites or areas transformed into coffee plantations. Around Turrialba and Juan Viñas observed and predicted areas correspond to sugar cane plantations. Some of these losses can be attributable to the harvest of the cane itself and the rest because still, some areas are suitable for this crop. Finally in the northeastern

part, at the east side of the Reventazón river, the loss seems to be related to the potential conversion of land to agroforestry systems or pastures, besides, slope and elevation have the lowest values for the basin in these areas.

Previous studies that have analyzed the additionality of established protected areas in avoiding deforestation had to search for counterfactuals. This is, finding areas with similar characteristics as the protected ones and check for differences in deforestation trends between them (Pfaff et al., 2014, 2009b). Previous studies have traditionally used a variety of matching methods to find counterfactuals and had to extract several variables to compare similar areas shown in tabular format (Andam et al., 2008; Ferraro, 2009; Pfaff et al., 2009b; Sanchez-Azofeifa et al., 2003). The RF model presented here is by itself a counterfactual, because it already divides the landscape into areas with similar characteristics (at-risk or not of vegetation loss), based only on physical variables. Additionally, it is spatially explicit, with different levels of probability, which may be helpful for decision-makers to complement existing frameworks for conservation targeting and to prioritize threatened areas.

Additional socioeconomic factors, like environmental legislation, rural exodus, decrease in forest pressure related to a decrease in meat exports, and the boost of ecotourism in recent years, are additional reasons that would explain the low rates of vegetation loss observed and predicted (Jadin et al., 2016; Stan and Sanchez-Azofeifa, 2019; Vallet et al., 2016). However, the model was only based on biophysical factors, and socioeconomic variables were not included because previous studies have shown little effect or importance of socioeconomic variables incorporated into RF models compared to physical and environmental conditions. Studies in tropical forests have shown that population density, for example, cannot be used directly as a proxy of deforestation, since there are complex mechanisms between population density and deforestation, with a variety of results not always leading to deforestation (Redo et al., 2012).

5.1.3.3. Historical forest loss in farms

Overall, regardless of whether the farms participated or not in the PES program, historical vegetation loss rates were very low. For all the forest areas found inside the studied farms, 94% of their forest remained undisturbed throughout the period 2000-2018. In general, deforestation was higher for farms not participating in the PES program and often related to the logging of tree plantations. Because during the period analyzed (2000-2018) the PES program was already enforced (started in 1997) these results might be seen as a

natural outcome of the program itself, indicating that the program had a positive effect on the decrease in forest loss. However, the results so far also indicate that forest lands included in the program would be biased towards lower pressure, due to the low opportunity cost, so that they would have remained forested even without receiving payments (Pfaff et al., 2008; Robalino and Pfaff, 2013; Sierra and Russman, 2006). Farms participating in the PES program showed vegetation losses also in small plantations or due to clearing needed for new infrastructures. These results would also indicate that the main drivers of deforestation are not related to the need for the increase in agricultural lands at a large scale but instead related to small adjustments around productive areas. Observed logging in plantations also can be interpreted as being an authorized activity, which would require the permission of authorities. Historical rates would be in line with previous research in the country, which found low deforestation rates and no substantial differences between participants and non-participants in the PES program in the initial years 1997-2000 (Pfaff et al., 2008; Sánchez-Azofeifa et al., 2007). Types E and F which show the largest share of forest area for both groups also showed the lowest vegetation losses.

5.1.3.4. Predicted risk of forest loss in farms

Based on the historical vegetation loss, the model predicts that forests will be lost on farms that belong to types where land use and land productivity are more important. As the characteristics of the farm allow the establishment of productive activities derived from land use, such as sugar cane, coffee, or agriculture, the risk of forest loss increases. Because the opportunity cost of leaving an area with protected forest can be higher compared with the transformation of this area to alternative uses, and even not compensated by the payment obtained from the PES program, these farms will correspond to non-PES farms which present higher risks of forest loss. They are located in areas where land-use change or tree removal is more frequent, due to forest converted to agriculture, pasture, or urbanization and are fairly accessible. On the contrary, farms participating in the PES program, show, in general, lower risks of forest loss. Very likely the opportunity cost of conserve forest areas is very low, because there is no alternative to the presence of forest due to low accessibility, high humidity, or cloudiness, and because the presence of forest must be positive for the objectives of the farm, often dedicated to conservation or tourism (Allen and Colson, 2019; Pfaff et al., 2008; Zbinden and Lee, 2005). These results would be supported by the finding that the predicted risk increases closer to land productive areas.

5.1.3.5. Evaluation of deforestation risk probability in PES and non-PES forested areas

Overall, the predicted deforestation risk on forested areas was low for all farms and categories, participating or not in the PES program. The analysis of 10 similar farms for each group in terms of share of forest size and total size, showed that farms participating in the PES program had larger forested areas (Arriagada et al., 2009) and lower deforestation risk than those not participating. In non-PES farms, higher risks were observed in small to medium-size farms, that correspond to those farms with intensive land use, where eventually tree clearing can be interesting from an economic perspective. Besides, PES farms seem to have a smaller proportion of productive land uses of the total land area than non-PES farms which may indicate a lower opportunity cost (Arriagada et al., 2009).

When further comparisons were made between the physical forest features of both groups, forests in PES farms seemed to be located in less accessible areas and further away from productive land uses (Mohebalian and Aguilar, 2018), which would support the hypotheses from previous authors that forests participating in the PES program would be under relatively low deforestation pressure (Fiorini et al., 2020; Robalino and Pfaff, 2013). Moreover, A potential outcome of the analysis of the different deforestation risks between PES and non-PES farms is that the program would be potentially biased to protect large forest areas at very low risk while overlooking potential forest losses in areas at higher risks.

These results do not question that the environmental services generated by forests are ultimately provided to their owners and society. It is proposed that the risk of forest loss, and therefore the decrease in the amount or continuity of these environmental services, would be spatially located under different conditions from where historically the program has been directing the payments.

5.2. Farmers' participation in PES: Landscape influences farming strategies

5.2.1. Farm typology, opportunity cost, and proportion of forest area

The creation of a typology was useful to understand the different management strategies adopted by the farmers and to understand their participation or not in the PES program. It also served to establish a solid idea on the composition of the landscape in the studied area and to understand the motivations that affect land use and how this can have an effect on the distribution of the forested areas inside the farm. Distribution of the different land uses inside a farm, is a combination of farmer decisions but is also biophysical. Biophysical characteristics, and especially topography, constraints the type of activities that can be developed from an agroecological and economic point of view. Besides, the analysis of farms based on field visits was useful to understand the peculiarities of the study site that for example allow the cultivation of crops at high slopes and elevations, contrary to what would be expected.

The distribution and location of trees and forests inside the farms are controlled by a combination of topography and opportunity cost. Where land is intensively used for agriculture, sugar cane, or pasture, there is little room for trees and they are frequently found in the least accessible areas, in shallow soils and steep slopes. Trees in productive areas are more abundant in coffee systems and increasingly being planted to shade coffee plants (Meylan et al., 2017). In pastures, scattered trees are beneficial to shade cattle and the growth of pasture is higher under the tree canopy. Moreover, they can be used as live fences and in some farms as fodder to feed cattle (Sibelet et al., 2017). Trees do not provide any economic benefit derived from their removal that can compete with agricultural activities and it was not found any interesting direct use from them other than making firewood from dead trees. Besides, any gain in tree cover, for example, if the forest were allowed to grow in a fallow area, it would entail a high opportunity cost for the owner, since once an area is considered forest, the 1996 Costarican forest law of 7575 would penalize its removal. Costa Rica's Forestry Law 7575 states that areas within 15 m of a watercourse and on slopes greater than 45 degrees are considered protected areas where logging is forbidden (Sibelet et al., 2017).

However, a forest can be an interesting source of income without the need for tree clearing, just conserving it in the most natural way possible. Tourism initiatives in farms benefit very much from this forest conservation, which is an interesting income alternative that requires very little or no intervention on the forest itself. Regarding opportunity cost, it is the lowest, agriculture is neither possible nor interesting in these farms, due to poor soils and difficult accessibility or slope, so the next best alternative would be to do nothing.
Therefore farmers would be only willing to include low-productivity land in the PES program and keep their high-productivity land for themselves (Sheng et al., 2019). Some respondents said that they would like to receive the payment for the area of forest that they cannot use for other uses. Also, because the participation in the program is voluntary, it is plausible to think that the farmer will select the least profitable and more marginal land to be included in the program (Robalino and Pfaff, 2013; Sánchez-Azofeifa et al., 2007; Zbinden and Lee, 2005). Previous studies in western Amazonia have shown that land price increases with increasing accessibility and presence of productive land use, such as agriculture, pasture, or coffee. However, the total area of farm, slope, and forest coverage are factors affecting negatively to land price (Holland et al., 2016). Also, the abandonment of difficult terrain can be interesting in favor of the booming tourism industry as it has been reported in other parts of Costa Rica (Allen and Vásquez, 2017).

At the farm level, the largest forest patches inside the farms increase as the total farm area increases, consistent with findings in other studies in the country (Zbinden and Lee, 2005). Contrary to agricultural areas, farms where the forest is predominant, are comparatively much larger. Over time, the agricultural lands that originate as a result of the transformation of the forest, if they are good for cultivation, are divided more and more, also increasing their price (Holland et al., 2016). Using cadastral maps of the area, it was also found that they tend to be located far from urban areas and in less accessible terrain. Therefore, concerning the participation in the PES program, it is important to highlight that there is a bias in the location of farms with large forest areas which, from an economic perspective, are potentially more interesting to be included in the program.
Farm distance from urban areas or main roads and the largest proportion of forest area are directly related to higher participation in PES programs in Costa Rica (Robalino and Pfaff, 2013), or Brazil (Fiorini et al., 2020)

5.2.2. Farmers' perception of forests and the PES program

All farmers interviewed in the area were aware of the importance of forest conservation and its benefits. In particular, the role of the forest as a water regulator was found the most important factor as well as in other parts of the country (Allen and Vásquez, 2017). Practices related to the conservation of biodiversity have been applied to the main crops in the area, like coffee, implementing agroforestry systems and tree shadowing, and trying to promote the use of organic products or biological pest control to avoid water pollution in the area. Concerning the conservation of trees, this is seen differently by farmers depending on the landscape where they are. In agricultural areas, farmers perceive that more trees are needed in the landscape, to improve environmental conditions, but they lack access and resources to good seedlings or plants and lack technical expertise (Sibelet et al., 2017).

The environmental problems found in the agricultural area, in particular, related to soil erosion were shocking and the PES program with its current modalities did not seem to offer a solution to this problem. However, the importance of taking action on these environmental problems is demonstrated by the economic spending that ICE makes each year in this area, among other things providing autochthonous trees to farmers (Vignola et al., 2010). In these areas, the conservation of the present trees remaining from a large-scale land conversion of the cloud forest that started in 1960, seemed more of a priority than the protection of vast forested cloudy areas close to Tapantí National Park. Farmers with small portions of forest in agricultural areas stated that the low payment received by the PES program and the restrictions imposed over the protected area, would not be interesting enough to the family economy.

About the perception of the PES program, the lack of clear information and mistrust in the administration staff seemed to be key to discourage the participation of some farmers. These trends have been maintained over time, as previous studies also found the same perceptions in Costa Rica (Arriagada et al., 2009; Zbinden and Lee, 2005) and seem to be frequent in similar PES programs in other countries, like in Brazil (Alarcon et al., 2017). The mistrust in the administration is reinforced by the perception from the farmer's perspective that the Government will impose many restrictions on their lands. This is discouraging for farmers with very intense use of the land, such as for agriculture, pasture, cane sugar, or coffee. These farmers may need to make changes in their land use if

needed due to economic reasons. The extreme case in the perception of Government control is reflected with testimonials of direct appropriation of farmer's land, highlighting the lack of information for some farmers. The interviews suggested that the level of education is also a key factor that determines participation. Owners that are used to deal with and understand administrative management issues are more in favor of participating (Alarcon et al., 2017; Zbinden and Lee, 2005).

Finally, it was found a potential relationship based on the costs associated with participation in the PES program i.e., the transaction costs and the size of the forest area. The amount directed to forestry managers needed to write a management plan and visit the farm can be costly, from about 10% of the received payment based on the interviews or up to 20-40% based on other studies (Allen and Colson, 2019; Wünscher et al., 2006). Because the conservation payment is received on a hectare basis, it is hypothesized that there is a certain threshold in the amount of forest area and consequently of payment that the owner must receive at least, to cover these costs and still that the remaining amount is interesting. Additionally, forest organizations that facilitate the administrative management related to these transaction costs might also be interested in handling larger farms. It seems obvious that owners with larger forest farms will have a preference to participate in the program (Zbinden and Lee, 2005).

5.3. The topography factor and its influence in the monitoring of forests and environmental services using remote sensing

5.3.1. Illumination conditions and vegetation indices

The IC model provided a detailed view of the effect of sun-sensor geometry and topography at a pixel level. The research reveals that frequently shadowed areas experience more variation in IC than sunlit areas. The increase in IC values with the improvement of sun-sensor geometry has been reported in other studies, but usually comparing two images at different times or several images across a season in the same year (Galvão et al., 2016; Matsushita et al., 2007; Wu et al., 2008). The effect of terrain in the correlations between EVI~IC and NDVI~IC has important implications when interpreting VI, especially EVI, which showed higher sensitivity. Because EVI~IC correlations were positive in shadowed areas and neutral to negative in sunlit areas, it can be interpreted that the areas with different IC experienced “greening” or “browning” when in reality there was no change in the vegetation conditions. The explanation for this would be that sunlit areas keep high IC values with low variation due to more constant sun-sensor geometry conditions. This would have a low effect on the variation of EVI and NDVI which correspond to neutral or even negative correlation values. However, shadowed areas experience a higher range of IC values which would have a higher effect on EVI and NDVI values, showing positive correlations. The strong effect observed in EVI in this irregular terrain corroborates the need for removing the topographic effect in the reflectance data before calculating this index (Matsushita et al., 2007). Therefore, the results using EVI for modeling phenology patterns, vegetation functioning or GPP would be biased if the IC effect is omitted, in accordance with research supporting this effect (Brede et al., 2015; Cao et al., 2016; Galvão et al., 2011; Saleska et al., 2016). In relation to the research that claims the effect of IC on VI, most of these studies used coarser resolution than Landsat and did not employ a DEM to simulate IC. In addition, most of them were carried out in the Amazon region in relatively flat terrain and yet IC effects were patent (Galvão et al., 2016, 2011; Maeda and Galvão, 2015; Morton et al., 2014). This agrees with previous studies that found drastic forest changes comparing various images in irregular terrain (Tan et al., 2013) and it is suggested that using IC is critical to understand these patterns. Although using a DEM with the same scale as the satellite data

improved the results of the interpretation, caution is required because of known issues in SRTM data such as shadows that can bias the results, further investigation is needed (Li et al., 2015). Despite this, the spatial resolution of the images used, allowed more detailed and better interpretations of changes in IC than sensors with coarser spatial resolution such as MODIS or methodologies relying upon angle information of the image (Galvão et al., 2016, 2011; Maeda and Galvão, 2015; Morton et al., 2014)

VI are also sensitive to vegetation structure factors such as LAI or canopy shadow which very likely affect the variations observed, but lower resolutions than those used in this study would be required, especially for EVI (Brede et al., 2015). Factors such as the background influence of soil and leaf litter or LAI are probably responsible for quickly saturating NDVI (Gao et al., 2000), hiding the effect of IC variation on this index.

5.3.2. Temporal analysis of illumination conditions

Temporal analysis for the whole Landsat reflectance collection in the area showed significant and increasing trends for IC, EVI, and NDVI. Information on zenith and azimuth solar angles already present in each available image was employed, eliminating the need for merging information between Landsat and MODIS sensors to model or simulate data. Intra- and inter-annual variations of VI as a response to changes in sun-sensor geometry and variations in sun-sensor geometry itself have been conducted using MODIS and Landsat collections, but covering shorter periods (Nagol et al., 2014; Zhang et al., 2016; Zhang and Roy, 2016). Some of the observed changes in IC agree with observations reported after analyzing the orbit change in Landsat 5 between the years 1995-2000, but here, IC variations are described better at the terrain level, something that was not possible to reflect in that study (Zhang and Roy, 2016). Although the terrain is believed to remain constant, the use of the SRTM dataset to model IC throughout the whole period could be a source of bias in some areas.

6. Conclusions

6.1. Predictive models of vegetation loss risk in protected areas

Developing reliable tools to predict land cover change and specifically, vegetation loss or deforestation is a priority in environmental sciences. It has been shown that RF provides high accuracy in predicting vegetation loss based on known free-access historical data. In the study site, accessibility predictor variables such as distance to main roads, elevation, or slope were the most important to predict vegetation loss. It is proposed that these are key factors that have to be considered by decision-makers when establishing protected areas. The model predicts that most of the vegetation loss will occur outside the PWA, with only a predicted loss of 4.43% of the initial vegetation inside the PWA in the year 2000.

These results question the effectiveness and additionality of keeping these areas protected and the potential need to create new tools of protection on areas where the predicted risk of deforestation is higher. These tools should include negotiations with local stakeholders to allow the coexistence of economic activities and forest protection. While previous studies have reached similar conclusions regarding the questionable effectiveness of the location of protected areas in Costa Rica, this is to the latest knowledge, the first study that is spatially explicit at the pixel level in illustrating the threat to which these areas are exposed. This research can serve as a basis for quantifying the change of different environmental services based on the predicted risk, and therefore serve to guide environmental political decisions in the allocation of resources to fight climate change.

6.2. The influence of topography on land use, opportunity cost, and participation in the PES program

This study reveals that the combination of qualitative and quantitative tools is a reliable methodology to understand forest conservation and the PES program in Costa Rica. Qualitative and quantitative analysis is needed because there are complex relationships between how topography drives land-use distribution inside the farms and how this limits economic activities, which could not be detected only using a single lens.
Observations and differences derived from biophysical variables from farms participating and non-participating in the PES program were confirmed by farmers' perceptions in the interviews.

Low accessibility and irregular topography drive the size of the farms in mountainous regions, increasing in size as these factors increase. Besides, these factors impose limits to the transformation of land use in these marginal areas into more productive and profitable activities. The result is that historically these areas have remained undisturbed and covered by forest. Also, economic activities in the area that depend on a good state of conservation of the forest, such as tourism, may prevent it from suffering any type of degradation or disturbance. It is therefore questioned whether the PES had any effect, not in the environmental services provided, but on protecting the forest in a different way from what would have happened otherwise. Although farms participating in the PES program showed lower historical and predicted deforestation, this might be related more to the low opportunity cost of their lands and accessibility than the effect of the program itself. Also, because the participation in the PES program is voluntary, farmers would be interested to include marginal forest lands from which they do not profit, and the program would be therefore biased to target remote forested areas with a low risk of deforestation. Future schemes of PES should include accessibility as a key factor if additionality and efficiency are to be improved in targeting forest areas at risk. In this regard, updated maps of road networks are essential. Also, access conditions to the program and costs should be revised to decrease the bias to targeting large at-low-risk forested farms.

The predictive model developed provides biophysical and spatially explicit information to understand the drivers of forest loss, and the locations where this is likely to occur, which can improve decisions taken when designing environmental policies.

Previous studies have empirically analyzed the effect of the PES program on forest cover using matching methods for the program's initial years. To the latest knowledge, this is the first study that uses machine learning techniques to predict forest loss based on observed historical deforestation throughout 18 years. Many of the hypotheses of previous empirical studies have been confirmed using this approach, and additionally, the method provides a measure of the importance of the variables that most affect deforestation, which is important when designing these schemes.

While the findings shown in this study may be representative of this particular region, and cannot be extrapolated to others, the methods employed have shown the importance of using an interdisciplinary approach to understand all factors affecting the location of forests.

6.3. The influence of topography in remote sensing and in the monitoring of forests and environmental services throughout the time

In this study, a novel approach has been developed to use Landsat temporal series to evaluate changes in the illumination conditions and vegetation indices in forested areas in irregular terrain. Google Earth Engine was used to analyze 397 images for 28 years (1984-2017) of the latest Landsat Surface Reflectance collections. The main conclusions of the work can be described as follows:

In summary, the illumination condition of Landsat scenes can be easily calculated using Landsat image information and an elevation model of the same resolution. Changes in illumination conditions and their relation to vegetation indices can be effectively evaluated at the pixel level without needing to use ancillary data from other sensors or simulating data.

The analysis at the pixel level showed correlations between illumination conditions and vegetation indices with a strong effect in EVI and less in NDVI. Moreover, positive correlations were found in shadowed areas, whereas neutral to negative correlations were found in sunlit areas, This is a crucial finding to interpret vegetation indices derived from

Landsat data in irregular terrain. In this regard, the analysis at the pixel level revealed patterns that could not be detected using only solar angle information.

The long-term analysis revealed that there is an increasing trend in illumination conditions, EVI, and NDVI associated with the improvement in the position in Landsat sensors over time. These trends calculated at selected images at varying seasons were significant when placed in chronological order. This effect cannot be overlooked when employing vegetation indices in the study of phenology patterns or above-ground net primary productivity. The incorporation of illumination conditions into time-series analysis according to this method can provide additional data to understand the behavior of vegetation indices in irregular terrain.

7. References

Aguilar-Amuchastegui, N., Riveros, J.C., Forrest, J.L., 2014. Identifying areas of deforestation risk for REDD+ using a species modeling tool. Carbon Balance Manag. 9, 1–10. https://doi.org/10.1186/s13021-014-0010-5

Aide, T.M., Clark, M.L., Grau, H.R., López-Carr, D., Levy, M.A., Redo, D., Bonilla-Moheno, M., Riner, G., Andrade-Núñez, M.J., Muñiz, M., 2013. Deforestation and Reforestation of Latin America and the Caribbean (2001-2010). Biotropica 45, 262–271. https://doi.org/10.1111/j.1744-7429.2012.00908.x

Alarcon, G.G., Fantini, A.C., Salvador, C.H., Farley, J., 2017. Additionality is in detail: Farmers' choices regarding payment for ecosystem services programs in the Atlantic forest, Brazil. J. Rural Stud. 54, 177–186. https://doi.org/10.1016/j.jrurstud.2017.06.008

Allen, K.E., Colson, G., 2019. Understanding PES from the ground up: a combined choice experiment and interview approach to understanding PES in Costa Rica. Sustain. Sci. 14, 391–404. https://doi.org/10.1007/s11625-018-00653-w

Allen, K.E., Vásquez, S.P., 2017. Forest cover, development, and sustainability in Costa Rica: Can one policy fit all? Land use policy 67, 212–221. https://doi.org/10.1016/j.landusepol.2017.05.008

Andam, K.S., Ferraro, P.J., Hanauer, M.M., 2013. The effects of protected area systems on ecosystem restoration: A quasi-experimental design to estimate the impact of Costa Rica's protected area system on forest regrowth. Conserv. Lett. 6, 317–323. https://doi.org/10.1111/conl.12004

Andam, K.S., Ferraro, P.J., Pfaff, A., Sanchez-Azofeifa, G.A., Robalino, J.A., 2008. Measuring the effectiveness of protected area networks in reducing deforestation. Proc. Natl. Acad. Sci. U. S. A. 105, 16089–16094. https://doi.org/10.1073/pnas.0800437105

Arriagada, R.A., Sills, E.O., Pattanayak, S.K., Ferraro, P.J., 2009. Combining qualitative and quantitative methods to evaluate participation in costa rica's program of payments for environmental services. J. Sustain. For. 28, 343–367. https://doi.org/10.1080/10549810802701192

Barnes, M.D., Glew, L., Wyborn, C., Craigie, I.D., 2018. Prevent perverse outcomes from global protected area policy. Nat. Ecol. Evol. 2, 759–762. https://doi.org/10.1038/s41559-018-0501-y

Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca, C., Picard, N., Sparrow, B., Abraham, E.M., Aloui, K., Atesoglu, A., Attore, F., Bassüllü, Ç., Bey, A., Garzuglia, M., García-Montero, L.G., Groot, N., Guerin, G., Laestadius, L., Lowe, A.J., Mamane, B., Marchi, G., Patterson, P., Rezende, M., Ricci, S., Salcedo, I., Diaz, A.S.-P., Stolle, F., Surappaeva, V., Castro, R., 2017. The extent of forest in dryland biomes. Science (80-.). 356, 635–638. https://doi.org/10.1126/science.aam6527

Bax, V., Francesconi, W., 2018. Environmental predictors of forest change: An analysis of natural predisposition to deforestation in the tropical Andes region, Peru. Appl. Geogr. 91, 99–110. https://doi.org/10.1016/j.apgeog.2018.01.002

Bera, B., Saha, S., Bhattacharjee, S., 2020. Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of Silabati watershed, India. Trees, For. People 2, 100034. https://doi.org/10.1016/j.tfp.2020.100034

Bernard, F., de Groot, R.S., Campos, J.J., 2009. Valuation of tropical forest services and mechanisms to finance their conservation and sustainable use: A case study of Tapantí National Park, Costa Rica. For. Policy Econ. 11, 174–183. https://doi.org/10.1016/j.forpol.2009.02.005

Börner, J., Baylis, K., Corbera, E., Ezzine-de-Blas, D., Honey-Rosés, J., Persson, U.M., Wunder, S., 2017. The Effectiveness of Payments for Environmental Services. World Dev. 96, 359–374. https://doi.org/10.1016/j.worlddev.2017.03.020

Brede, B., Suomalainen, J., Bartholomeus, H., Herold, M., Brede, B., Suomalainen, J., Bartholomeus, H., Herold, M., 2015. Influence of solar zenith angle on the enhanced vegetation index of a Guyanese rainforest. Remote Sens. Lett. 7058. https://doi.org/10.1080/2150704X.2015.1089362

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/https://doi.org/10.1023/A:1010933404324

Cao, S., Sánchez-Azofeifa, G., Duran, S., Calvo-Rodriguez, S., 2016. Estimation of aboveground net primary productivity in secondary tropical dry forests using the Carnegie – Ames – Stanford approach (CASA) model. Environ. Res. Lett. 11. https://doi.org/10.1088/1748-9326/11/7/075004

Chandrasekar, K., Sesha Sai, M.V.R., Roy, P.S., Dwevedi, R.S., 2010. Land Surface Water Index (LSWI) response to rainfall and NDVI using the MODIS vegetation index product. Int. J. Remote Sens. 31, 3987–4005. https://doi.org/10.1080/01431160802575653

Chomitz, K.M., Alger, K., Thomas, T.S., Orlando, H., Nova, P.V., 2005. Opportunity costs of conservation in a biodiversity hotspot: The case of southern Bahia. Environ. Dev. Econ. 10, 293–312. https://doi.org/10.1017/S1355770X05002081

Cord, A.F., Brauman, K.A., Chaplin-Kramer, R., Huth, A., Ziv, G., Seppelt, R., 2017. Priorities to Advance Monitoring of Ecosystem Services Using Earth Observation. Trends Ecol. Evol. 32, 416–428. https://doi.org/10.1016/j.tree.2017.03.003

Cushman, S.A., Macdonald, E.A., Landguth, E.L., Malhi, Y., Macdonald, D.W., 2017. Multiple-scale prediction of forest loss risk across Borneo. Landsc. Ecol. 32, 1581–1598. https://doi.org/10.1007/s10980-017-0520-0

Daniels, A.E., Bagstad, K., Esposito, V., Moulaert, A., Rodriguez, C.M., 2010. Understanding the impacts of Costa Rica's PES: Are we asking the right questions? Ecol. Econ. 69, 2116–2126. https://doi.org/10.1016/j.ecolecon.2010.06.011

De Alban, J.D.T., Connette, G.M., Oswald, P., Webb, E.L., 2018. Combined Landsat and L-band SAR data improves land cover classification and change detection in dynamic tropical landscapes. Remote Sens. 10. https://doi.org/10.3390/rs10020306

De Araujo Barbosa, C.C., Atkinson, P.M., Dearing, J.A., 2015. Remote sensing of ecosystem services: A systematic review. Ecol. Indic. 52, 430–443. https://doi.org/10.1016/j.ecolind.2015.01.007

Di Lallo, G., Mundhenk, P., López, S.E.Z., Marchetti, M., Köhl, M., 2017. REDD+: Quick assessment of deforestation risk based on available data. Forests 8, 1–16. https://doi.org/10.3390/f8010029

Didan, K., Munoz, A.B., Solano, R., Huete, A., 2015. MODIS vegetation index User 's Guide (MOD13 Series). The University of Arizona: Tucson, AZ, USA.

Dlamini, W.M., 2016. Analysis of deforestation patterns and drivers in Swaziland using efficient Bayesian multivariate classifiers. Model. Earth Syst. Environ. 2, 1–14. https://doi.org/10.1007/s40808-016-0231-6

Drew, C.A., Wiersma, Y.F., Huettmann, F., 2010. Predictive Species and Habitat Modeling in Landscape Ecology: concepts and applications. Springer Science & Business Media.

Evans, J.S., Cushman, S.A., 2009. Gradient modeling of conifer species using random forests. Landsc. Ecol. 24, 673–683. https://doi.org/10.1007/s10980-009-9341-0

FAO, 2020. Global Forest Resources Assessment 2020: Main report, FAO. Rome. https://doi.org/https://doi.org/10.4060/ca9825en

FAO, 1990. Global Forest Resources Assessment 1990. Tropical countries, FAO. Rome.

FAO, UNEP, 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. Rev. Geophys 45, 1–33. https://doi.org/10.1029/2005RG000183.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Ferraro, P.J., 2009. Counterfactual thinking and impact evaluation in environmental policy. New Dir. Eval. 122, 75–84. https://doi.org/10.1002/ev.297

Ferraro, P.J., Kiss, A., 2012. Direct Payments to Conserve Biodiversity. Science (80-.). 298, 1718–1719.

Fiorini, A.C.O., Mullally, C., Swisher, M., Putz, F.E., 2020. Forest cover effects of payments for ecosystem services: Evidence from an impact evaluation in Brazil. Ecol. Econ. 169, 106522. https://doi.org/10.1016/j.ecolecon.2019.106522

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. Remote Sens. Environ. 194, 379–390. https://doi.org/10.1016/j.rse.2017.03.026

Galvão, L.S., Breunig, F.M., Teles, T.S., Gaida, W., Balbinot, R., 2016. Investigation of terrain illumination effects on vegetation indices and VI-derived phenological metrics in subtropical deciduous forests. GIScience Remote Sens. 53, 360–381. https://doi.org/10.1080/15481603.2015.1134140

Galvão, L.S., dos Santos, J.R., Roberts, D.A., Breunig, F.M., Toomey, M., de Moura, Y.M., 2011. On intra-annual EVI variability in the dry season of tropical forest: A case study with MODIS and hyperspectral data. Remote Sens. Environ. 115, 2350–2359. https://doi.org/10.1016/j.rse.2011.04.035

Gao, B.C., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gao, X., Huete, A.R., Ni, W., Miura, T., 2000. Optical-biophysical relationships of vegetation spectra without background contamination. Remote Sens. Environ. 74, 609–620. https://doi.org/10.1016/S0034-4257(00)00150-4

Geist, H.J., Lambin, E.F., 2002. Proximate causes and underlying driving forces of tropical deforestation. Bioscience 52, 143–150. https://doi.org/10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2

Goldman, E., Harris, N., Maschler, T., 2017. Predicting Future Forest Loss in the Democratic Republic of the Congo ' S Carpe Landscapes. Washington, D.C.

Gómez-Baggethun, E., de Groot, R., Lomas, P.L., Montes, C., 2010. The history of ecosystem services in economic theory and practice: From early notions to markets and payment schemes. Ecol. Econ. 69, 1209–1218. https://doi.org/10.1016/j.ecolecon.2009.11.007

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine : Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2017.06.031

Hansen, M.C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S.J., Loveland, T.R., Kommaredy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science (80-.). 342, 850–854. https://doi.org/10.1126/science.1244693

Hart, E., Sim, K., Kamimura, K., Meredieu, C., Guyon, D., Gardiner, B., 2019. Use of machine learning techniques to model wind damage to forests. Agric. For. Meteorol. 265, 16–29. https://doi.org/10.1016/j.agrformet.2018.10.022

Havinga, I., Hein, L., Vega-Araya, M., Languillaume, A., 2020. Spatial quantification to examine the effectiveness of payments for ecosystem services: A case study of Costa Rica's Pago de Servicios Ambientales. Ecol. Indic. 108, 105766. https://doi.org/10.1016/j.ecolind.2019.105766

Hethcoat, M., Edwards, D., Carreiras, J., Bryant, R., França, F., Quegan, S., 2019. A machine learning approach to map tropical selective logging. Remote Sens. Environ. 221, 1–6. https://doi.org/10.1101/451856

Holben, B.N., Justice, C., 1980. The topographic effect on spectral response from nadir-pointing sensors. Photogramm. Eng. Remote Sens. 46, 1191–1200.

Holland, T.G., Coomes, O.T., Robinson, B.E., 2016. Evolving frontier land markets and the opportunity cost of sparing forests in western Amazonia. Land use policy 58, 456–471. https://doi.org/10.1016/j.landusepol.2016.08.015

Hollander, M., Wolfe, D.A., Chicken, E., 2014. Nonparametric statistical methods, third edit. ed, Nonparametric Statistical Methods. John Wiley & Sons, Hoboken, New Jersey. https://doi.org/10.1002/9781119196037

Huete, A.R., 1988. A Soil-Adjusted Vegetation Index (SAVI). Remote Sens. Environ. 25, 295–309.

Huete, A.R., Didan, K., Shimabukuro, Y.E., Ratana, P., Saleska, S.R., Hutyra, L.R., Yang, W., Nemani, R.R., Myneni, R., 2006. Amazon rainforests green-up with sunlight in dry season. Geophys. Res. Lett. 33, 2–5. https://doi.org/10.1029/2005GL025583

Humphries, G.R., Magness, D.R., Huettmann, F., 2018. Machine Learning for Ecology and Sustainable Natural Resource Management, Machine learning for ecology and sustainable natural resource management. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-96978-7

IPCC, 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen,. Cambridge Univ. Press 3949.

Jadin, I., Meyfroidt, P., Lambin, E.F., 2016. International trade, and land use intensification and spatial reorganization explain Costa Rica's forest transition. Environ. Res. Lett. 11. https://doi.org/10.1088/1748-9326/11/3/035005

Joppa, L.N., Pfaff, A., 2010. High and Far : Biases in the Location of Protected Areas 4, 1–6. https://doi.org/10.1371/journal.pone.0008273

Karsenty, A., Ezzine-de-blas, D., 2016. PES, markets and property rights: a comment on Wunder's revisited concept of PES and a proposal of conceptual framework. <hal-01262380> 16.

Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J. V., Grainger, A., Lindquist, E., 2015. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. For. Ecol. Manage. 352, 9–20. https://doi.org/10.1016/j.foreco.2015.06.014

Leblois, A., Damette, O., Wolfersberger, J., 2017. What has Driven Deforestation in Developing Countries Since the 2000s? Evidence from New Remote-Sensing Data. World Dev. 92, 82–102. https://doi.org/10.1016/j.worlddev.2016.11.012

Li, F., Jupp, D.L.B., Thankappan, M., 2015. Issues in the application of Digital Surface Model data to correct the terrain illumination effects in Landsat images. Int. J. Digit. Earth 8, 235–257. https://doi.org/10.1080/17538947.2013.866701

Liu, H.Q., Huete, A., 1995. Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Trans. Geosci. Remote Sens. 33, 457–465. https://doi.org/10.1109/36.377946

Liu, K., Wang, S., Li, X., Li, Y., Zhang, B., Zhai, R., 2020. The assessment of different vegetation indices for spatial disaggregating of thermal imagery over the humid agricultural region. Int. J. Remote Sens. 41, 1907–1926. https://doi.org/10.1080/01431161.2019.1677969

Lovejoy, T.E., 2006. Protected areas: a prism for a changing world. Trends Ecol. Evol. 21, 329–333. https://doi.org/10.1016/j.tree.2006.04.005

Maeda, E.E., Galvão, L.S., 2015. Sun-sensor geometry effects on vegetation index anomalies in the Amazon rainforest. GIScience Remote Sens. 52, 332–343. https://doi.org/10.1080/15481603.2015.1038428

Maeda, E.E., Heiskanen, J., Aragão, L.E.O.C., Rinne, J., 2014. Can MODIS EVI monitor ecosystem productivity in the Amazon rainforest? Geophys. Res. Lett. 41, 7176–7183. https://doi.org/10.1002/2014GL061535

Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Watson, M.C., Amer, S., Weltz, M., Goodrich, D., Marsett, R., 2006. Remote sensing for grassland management in the arid Southwest. Rangel. Ecol. Manag. 59, 530–540. https://doi.org/10.2111/05-201R.1

Martínez, A., 2014. Evaluación fitosanitaria del vivero forestal Cachí del Instituto Costarricense de Electricidad (ICE) presente en la cuenca media del río Reventazón. Instituto Tecnológico de Costa Rica.

Mascaro, J., Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Martin, R.E., Anderson, C., Higgins, M., Chadwick, K.D., 2014. A tale of two "Forests": Random Forest machine learning aids tropical Forest carbon mapping. PLoS One 9, 12–16. https://doi.org/10.1371/journal.pone.0085993

Matsushita, B., Yang, W., Chen, J., Onda, Y., Qiu, G., 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: A case study in high-density cypress forest. Sensors 7, 2636–2651. https://doi.org/10.3390/s7112636

Mayfield, H., Smith, C., Gallagher, M., Hockings, M., 2017. Use of freely available datasets and machine learning methods in predicting deforestation. Environ. Model. Softw. 87, 17–28. https://doi.org/10.1016/j.envsoft.2016.10.006

Mellor, A., Haywood, A., Stone, C., Jones, S., 2013. The performance of random forests in an operational setting for large area sclerophyll forest classification. Remote Sens. 5, 2838–2856. https://doi.org/10.3390/rs5062838

Merry, F., Amacher, G., Lima, E., 2008. Land Values in Frontier Settlements of the Brazilian Amazon. World Dev. 36, 2390–2401. https://doi.org/10.1016/j.worlddev.2007.11.014

Meyfroidt, P., Rudel, T.K., Lambin, E.F., 2010. Forest transitions, trade, and the global displacement of land use. Proc. Natl. Acad. Sci. 107, 20917–20922. https://doi.org/10.1073/pnas.1014773107

Meylan, L., Gary, C., Allinne, C., Ortiz, J., Jackson, L., Rapidel, B., 2017. Evaluating the effect of shade trees on provision of ecosystem services in intensively managed coffee plantations. Agric. Ecosyst. Environ. 245, 32–42. https://doi.org/10.1016/j.agee.2017.05.005

Mohebalian, P.M., Aguilar, F.X., 2018. Design of tropical forest conservation contracts considering risk of deforestation. Land use policy 70, 451–462. https://doi.org/10.1016/j.landusepol.2017.11.008

Mokondoko, P., Manson, R.H., Ricketts, T.H., Geissert, D., 2018. Spatial analysis of ecosystem service relationships to improve targeting of payments for hydrological services. PLoS One 13, 1–27. https://doi.org/10.1371/journal.pone.0192560

Morse, W.C., Schedlbauer, J.L., Sesnie, S.E., Finegan, B., Harvey, C.A., Hollenhorst, S.J., Kavanagh, K.L., Stoian, D., Wulfhorst, J.D., 2009. Consequences of Environmental Service Payments for Forest Retention and Recruitment in a Costa Rican Biological Corridor 14.

Morton, D.C., Nagol, J., Carabajal, C.C., Rosette, J., Palace, M., Cook, B.D., Vermote, E.F., Harding, D.J., North, P.R.J., 2014. Amazon forests maintain consistent canopy structure and greenness during the dry season. Nature 506, 221–224. https://doi.org/10.1038/nature13006

Muradian, R., Arsel, M., Pellegrini, L., Adaman, F., Aguilar, B., Agarwal, B., Corbera, E., Ezzine de Blas, D., Farley, J., Froger, G., Garcia-Frapolli, E., Gómez-Baggethun, E., Gowdy, J., Kosoy, N., Le Coq, J.F., Leroy, P., May, P., Méral, P., Mibielli, P., Norgaard, R., Ozkaynak, B., Pascual, U., Pengue, W., Perez, M., Pesche, D., Pirard, R., Ramos-Martin, J., Rival, L., Saenz, F., Van Hecken, G., Vatn, A., Vira, B., Urama, K., 2013. Payments for ecosystem services and the fatal attraction of win-win solutions. Conserv. Lett. 6, 274–279. https://doi.org/10.1111/j.1755-263X.2012.00309.x

Nagol, J.R., Sexton, J.O., Kim, D.H., Anand, A., Morton, D., Vermote, E., Townshend, J.R., 2014. Bidirectional effects in Landsat reflectance estimates: Is there a problem to solve? ISPRS J. Photogramm. Remote Sens. 103, 129–135. https://doi.org/10.1016/j.isprsjprs.2014.09.006

Peng, D., Zhang, H., Yu, L., Wu, M., Wang, F., Huang, W., Liu, L., Sun, R., Li, C., Wang, D., Xu, F., 2018. Assessing spectral indices to estimate the fraction of photosynthetically active radiation absorbed by the vegetation canopy. Int. J. Remote Sens. 39, 8022–8040. https://doi.org/10.1080/01431161.2018.1479795

Pfaff, A., Robalino, J., Herrera, D., Sandoval, C., 2015. Protected areas' impacts on Brazilian Amazon deforestation: Examining conservation - Development interactions to inform planning. PLoS One 10, 1–17. https://doi.org/10.1371/journal.pone.0129460

Pfaff, A., Robalino, J., Lima, E., Sandoval, C., Herrera, L.D., 2014. Governance, Location and Avoided Deforestation from Protected Areas: Greater Restrictions Can Have Lower Impact, Due to Differences in Location. World Dev. 55, 7–20. https://doi.org/10.1016/j.worlddev.2013.01.011

Pfaff, A., Robalino, J., Sanchez-Azofeifa, A., Andam, K.S., Ferraro, P.J., 2009a. Location affects protection & the generation of REDD: Observable characteristics drive park impacts in Costa Rica. IOP Conf. Ser. Earth Environ. Sci. 6, 252023. https://doi.org/10.1088/1755-1307/6/25/252023

Pfaff, A., Robalino, J., Sanchez-azofeifa, G.A., Andam, K.S., Ferraro, P.J., 2009b. Park Location Affects Forest Protection : Land Characteristics Cause Differences in Park Impacts across Costa Rica. B.E. J. Econ. Anal. Policy 9, 1–24. https://doi.org/10.2202/1935-1682.1990

Pfaff, A., Robalino, J.A., Sanchez-Azofeifa, G.A., 2008. Payments for Environmental Services: empirical analysis for Costa Rica. Durham, NC, USA.

Phillips, A., 2004. The history of the international system of protected area management categories. Parks 14, 4–14.

Ponzoni, F.J., da Silva, C.B., dos Santos, S.B., Montanher, O.C., dos Santos, T.B., 2014. Local illumination influence on vegetation indices and plant area index (PAI) relationships. Remote Sens. 6, 6266–6282. https://doi.org/10.3390/rs6076266

Ponzoni, F.J., Galvão, L.S., Liesenberg, V., Santos, J.R., 2010. Impact of multi-angular CHRIS/PROBA data on their empirical relationships with tropical forest biomass. Int. J. Remote Sens. 31, 5257–5273. https://doi.org/10.1080/01431160903303005

Porras, I., Asquith, N., 2018. Ecosystems, poverty alleviation and conditional transfers. Guidance for practitioners.

Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems 9, 181–199. https://doi.org/10.1007/s10021-005-0054-1

Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 9, 1–19. https://doi.org/10.1002/widm.1301

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Redo, D.J., Aide, T.M., Clark, M.L., 2012. The Relative Importance of Socioeconomic and Environmental Variables in Explaining Land Change in Bolivia, 2001-2010. Ann. Assoc. Am. Geogr. 102, 778–807. https://doi.org/10.1080/00045608.2012.678036

Robalino, J., Pfaff, A., 2013. Ecopayments and deforestation in Costa Rica: A nationwide analysis of PSA's initial years. Land Econ. 89, 432–448. https://doi.org/10.3368/le.89.3.432

Robalino, J., Pfaff, A., Villalobos, L., 2017. Heterogeneous local spillovers from protected areas in Costa Rica. J. Assoc. Environ. Resour. Econ. 4, 795–820. https://doi.org/10.1086/692089

Robalino, J., Pfaff, A., Villalobos, L., 2011. Assessing the Impact of Institutional Design of Payments for Environmental Services: The Costa Rican Experience. Ecosyst. Serv. from Agric. AgroforestryMeasurement Paym. 305–318.

Rouse, J.W., Hass, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. Third Earth Resour. Technol. Satell. Symp. 1, 309–317. https://doi.org/citeulike-article-id:12009708

Sader, S.A., Joyce, A.T., 1988. Deforestation Rates and Trends in Costa Rica, 1940 to 1983. Biotropica 20, 11. https://doi.org/10.2307/2388421

Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P.T.T., Paul, G.C., 2020. Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India. Sci. Total Environ. 730, 139197. https://doi.org/10.1016/j.scitotenv.2020.139197

Saleska, S.R., Wu, J., Guan, K., Araujo, A.C., Huete, A., Nobre, A.D., Restrepo-Coupe, N., 2016. Dry-season greening of Amazon forests. Nature 531, E4–E5. https://doi.org/10.1038/nature16457

Salzman, J., Bennett, G., Carroll, N., Goldstein, A., Jenkins, M., 2018. The global status and trends of Payments for Ecosystem Services. Nat. Sustain. 1, 136–144. https://doi.org/10.1038/s41893-018-0033-0

Sanchez-Azofeifa, G.A., Daily, G.C., Pfaff, A.S.P., Busch, C., 2003. Integrity and isolation of Costa Rica 's national parks and biological reserves:examining the dynamics of land-cover change. Biol. Conserv. 109, 123 –135.

Sánchez-Azofeifa, G.A., Pfaff, A., Robalino, J., Boomhower, J.P., 2007. Costa Rica's payment for environmental services program: Intention, implementation, and impact. Conserv. Biol. 21, 1165–1173. https://doi.org/10.1111/j.1523-1739.2007.00751.x

Sánchez, O., Navarrete, G., 2017. The Experience of Costa Rica with the Payments for Environmental Services: 20 Years of Lessons Learned. Trop. J. Environ. Sci. 51, 195–214.

Sheng, J., Qiu, H., Zhang, S., 2019. Opportunity cost, income structure, and energy structure for landholders participating in payments for ecosystem services: Evidence from Wolong National Nature Reserve, China. World Dev. 117, 230–238. https://doi.org/10.1016/j.worlddev.2019.01.016

Shirvani, Z., 2020. A holistic analysis for landslide susceptibility mapping applying geographic object-based random forest: A comparison between protected and non-protected forests. Remote Sens. 12, 1–22. https://doi.org/10.3390/rs12030434

Sibelet, N., Mutel, M., Arragon, P., Luye, M. 2013. Qualitative survey methods applied to natural resource management. Online learning modules. Available at: http://entretiens.iamm.fr/

Sibelet, N., Chamayou, L., Newing, H., Montes, I.G., 2017. Perceptions of Trees Outside Forests in Cattle Pastures: Land Sharing Within the Central Volcanic Talamanca Biological Corridor, Costa Rica. Hum. Ecol. 45, 499–511. https://doi.org/10.1007/s10745-017-9924-3

Sibelet, N., Posada, K.E., Gutiérrez-Montes, I.A., 2019. Agroforestry systems provide firewood for livelihood improvement in Guatemala. Bois Forets des Trop. 340, 91–102. https://doi.org/10.19182/bft2019.340.a31692

Sierra, R., Russman, E., 2006. On the efficiency of environmental service payments: A forest conservation assessment in the Osa Peninsula, Costa Rica. Ecol. Econ. 59, 131–141. https://doi.org/10.1016/j.ecolecon.2005.10.010

Sills, E.O., Caviglia-Harris, J.L., 2009. Evolution of the Amazonian frontier: Land values in Rondônia, Brazil. Land use policy 26, 55–67. https://doi.org/10.1016/j.landusepol.2007.12.002

Silva, A.C.O., Fonseca, L.M.G., Körting, T.S., Escada, M.I.S., 2020. A spatio-temporal Bayesian Network approach for deforestation prediction in an Amazon rainforest expansion frontier. Spat. Stat. 35, 100393. https://doi.org/10.1016/j.spasta.2019.100393

Smith, J.A., Lie Lin, T., Ranson, K.J., 1980. The Lambertian assumption and Landsat data. Photogramm. Eng. Remote Sens. 46, 1183–1189.

Stan, K., Sanchez-Azofeifa, A., 2019. Deforestation and secondary growth in Costa Rica along the path of development. Reg. Environ. Chang. 19, 587–597. https://doi.org/10.1007/s10113-018-1432-5

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8. https://doi.org/10.1186/1471-2105-8-25

Tan, B., Masek, J.G., Wolfe, R., Gao, F., Huang, C., Vermote, E.F., Sexton, J.O., Ederer, G., 2013. Improved forest change detection with terrain illumination corrected Landsat images. Remote Sens. Environ. 136, 469–483. https://doi.org/10.1016/j.rse.2013.05.013

Teillet, P.M., Guindon, B., Goodenough, D.G., 1982. On the slope-aspect correction of multispectral scanner data. Can. J. Remote Sens. 8, 84–106. https://doi.org/10.1080/07038992.1982.10855028

USGS, 2019a. Landsat 4-7 Surface Reflectance (Ledaps) Product Guide, USGS. EROS: Sioux Falls, SD, USA. https://doi.org/10.1016/0042-207X(74)93024-3

USGS, 2019b. Landsat 8 Surface Reflectance Code (LASRC) Product Guide. EROS: Sioux Falls, SD, USA.

USGS, 2019c. Landsat Collection 1 Level 1 Product Definition. EROS: Sioux Falls, SD, USA.

Valle, D., Hyde, J., Marsik, M., Perz, S., 2020. Improved inference and prediction for imbalanced binary big data using case-control sampling: A case study on deforestation in the Amazon region. Remote Sens. 12. https://doi.org/10.3390/RS12081268

Vallet, A., Locatelli, B., Levrel, H., Pérez, C.B., Imbach, P., Carmona, N.E., Manlay, R., Oszwald, J., 2016. Dynamics of ecosystem services during forest transitions in Reventazón, Costa Rica. PLoS One 11, 1–18. https://doi.org/10.1371/journal.pone.0158615

Van Den Hoek, J., Smith, A.C., Hurni, K., Saksena, S., Fox, J., 2021. Shedding new light on mountainous forest growth: A cross-scale evaluation of the effects of topographic illumination correction on 25 years of forest cover change across nepal. Remote Sens. 13. https://doi.org/10.3390/rs13112131

Van Deventer, A.P., Ward, A.D., Gowda, P.M., Lyon, J.G., 1997. Using thematic mapper data to identify contrasting soil plains and tillage practices. Photogramm. Eng. Remote Sensing 63, 87–93.

Vignola, R., Koellner, T., Scholz, R.W., McDaniels, T.L., 2010. Decision-making by farmers regarding ecosystem services: Factors affecting soil conservation efforts in Costa Rica. Land use policy 27, 1132–1142. https://doi.org/10.1016/j.landusepol.2010.03.003

Willcock, S., Martínez-López, J., Hooftman, D.A.P., Bagstad, K.J., Balbi, S., Marzo, A., Prato, C., Sciandrello, S., Signorello, G., Voigt, B., Villa, F., Bullock, J.M., Athanasiadis, I.N., 2018. Machine learning for ecosystem services. Ecosyst. Serv. 33, 165–174. https://doi.org/10.1016/j.ecoser.2018.04.004

Wolf, C., Levi, T., Ripple, W.J., Zárrate-Charry, D.A., Betts, M.G., 2021. A forest loss report card for the world's protected areas. Nat. Ecol. Evol. 5, 520–529. https://doi.org/10.1038/s41559-021-01389-0

Wu, J., Bauer, M.E., Wang, D., Manson, S.M., 2008. A comparison of illumination geometry-based methods for topographic correction of QuickBird images of an undulant area. ISPRS J. Photogramm. Remote Sens. 63, 223–236. https://doi.org/10.1016/j.isprsjprs.2007.08.004

Wunder, S., 2007. The efficiency of payments for environmental services in tropical conservation: Essays. Conserv. Biol. 21, 48–58. https://doi.org/10.1111/j.1523-1739.2006.00559.x

Wünscher, T., Engel, S., Wunder, S., 2006. Payments for environmental services in Costa Rica: Increasing efficiency through spatial differentiation. Q. J. Int. Agric. 45, 319–337. https://doi.org/10.1016/j.ecolecon.2007.07.033

Zanella, L., Folkard, A.M., Blackburn, G.A., Carvalho, L.M.T., 2017. How well does random forest analysis model deforestation and forest fragmentation in the Brazilian Atlantic forest? Environ. Ecol. Stat. 24, 529–549. https://doi.org/10.1007/s10651-017-0389-8

Zar, J.H., 2014. Biostatistical Analysis, 5th ed, Prentice Hall. Prentice Hall, Upper Saddle River, New Jersey. https://doi.org/10.1142/9789814566780_0009

Zbinden, S., Lee, D.R., 2005. Paying for Environmental Services: An analysis of participation in Costa Rica's PSA program. World Dev. 33, 255–272. https://doi.org/10.1016/j.worlddev.2004.07.012

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 Sens. Environ. 186, 217–233. https://doi.org/10.1016/j.rse.2016.08.022

Zhang, H.K., Roy, D.P., Kovalskyy, V., 2016. Optimal solar geometry definition for global long-term landsat time-series bidirectional reflectance normalization. IEEE Trans. Geosci. Remote Sens. 56, 3624. https://doi.org/10.1109/TGRS.2018.2827478

02-UPM
(2021)

Modelling PES (Payments for Ecosystem Services) to prevent deforestation in tropical montane cloud forests (TMCFs) and meet the socioeconomic needs of their farmers in Latin America

Pablo Martín Ortega