UNIVERSITY OF HELSINKI
FACULTY OF AGRICULTURE AND FORESTRY

# Assessing protected area effectiveness in Madagascar using three satellite-based datasets – A study using matching methods

Saija Papunen
Master's Thesis
University of Helsinki
Master's Programme of Forest Sciences
Forest Ecology and Management
April 2022

| Tiedekunta/Fakultet – Faculty<br>Faculty of Agriculture and Forestry | Osasto/Institution– Department<br>The Department of Forest Sciences | |
|---|---|---|
| Tekijä/Författare – Author<br>Saija Papunen | | |
| Työn nimi / Arbetets titel – Title<br>Assessing protected area effectiveness in Madagascar using three satellite-based datasets – A study using matching methods | | |
| Oppiaine /Läroämne – Subject<br>Forest Ecology and Management | | |
| Työn laji/Arbetets art – Level<br>Pro gradu | Aika/Datum – Month and year<br>April 2022 | Sivumäärä/ Sidoantal – Number of pages<br>53 |

Tiivistelmä/Referat – Abstract

Measuring the effectiveness of protected areas (PAs) is essential as they are key tools in tackling the ongoing biodiversity loss and there is substantial variation in their effectiveness (the estimated ability of protected areas to prevent unnatural disturbances). In forested PAs, the most common variable in effectiveness estimation is forest loss, but fire can also be used as a proxy for conversion. There is, however, a lack of robust comparisons between different data sets and proxies. This thesis aims to provide more insight into the issue by comparing three satellite-based data sets in protected area effectiveness assessment using Madagascar as a case study. The questions to be answered here are whether the data sets and variables derived from them produce similar PA effectiveness estimates and whether they could be used interchangeably in research and for practical management purposes. The hypotheses are as follows: H1: The three proxies produce similar results with the two fire proxies most likely having a stronger relationship. H2: The data sets can be used interchangeably both for science purposes and in practical management of PAs.

The effectiveness of Malagasy protected areas established in or before 2005 (N=42) was examined from 2005 to 2017. Three binary response variables were compared: forest loss, fire incidence, and burned area. In addition, a continuous forest loss variable was examined. Forested areas and the full landscape were studied separately i.e. estimates were produced for both forested areas only and full landscape (forested areas + other areas). 1-kilometre parcels in a uniform grid were sampled using nearest neighbour Mahalanobis distance matching, controlling for the factors affecting conversion pressures with appropriate covariates: altitude, slope, distance to cities, distance to roads, distance to waterways, and rainfall for forested areas and full landscape, and in addition, distance to forest edge for forested areas. Relative effect, pooled relative effect, and network relative effect were calculated for the binary variables, mean effect for the continuous variable. The effects were calculated on country level, biome level (tropical and subtropical moist broadleaved forests, tropical and subtropical dry broadleaved forests, and deserts and xeric shrublands), and individual PA level.

Protected areas appeared to be at least moderately effective, and all variables produced parallel, consistent results on the country and biome level, especially when using pooled relative effect. On average, PAs in tropical and subtropical moist broadleaf forests were most effective in avoiding land-use pressures, the ones in tropical and subtropical dry broadleaf forests slightly less, and the ones in deserts and xeric shrublands most ineffective. There was substantial variation between and inside individual PAs, and in approximately half of the PAs all variables indicate that the given area is significantly effective ($\alpha = 0,05$). In a little over half of the PAs the effects were mixed, and in forested areas, no PA was indicated to be ineffective by all variables. In full landscape, this was the case for one PA. There were small differences between forested areas and the full landscape in all levels, but they were statistically significant only in a few cases.

This study thus suggests that the data sets could be used interchangeably, at least on country and biome level, when conducting matching to assess PA effectiveness in a tropical setting. They could be utilised on individual PA level, too, with certain precautions and understanding of the nature and behaviour of the data. They are well suited for research; however, in practical management forest loss and fire incidence might be more feasible than burned area, due to its certain characteristics (it for example demands quite a lot of processing depending on the use purpose) and accessibility issues.

Avainsanat – Nyckelord – Keywords
Madagascar, protected areas, protected area effectiveness, matching, Mahalanobis distance

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Prof. Markus Holopainen, Dr. Johanna Eklund

Säilytyspaikka – Förvaringställe – Where deposited
Helsinki University Library – Helda / E-thesis (opinnäytteet) ethesis.helsinki.fi

Muita tietoja – Övriga uppgifter – Additional information

| Tiedekunta/Fakultet – Faculty<br>Maatalous-metsätieteellinen tiedekunta | Osasto/Institution– Department<br>Metsätieteiden osasto |
|---|---|


Tekijä/Författare – Author
Saija Papunen


Työn nimi / Arbetets titel – Title
Assessing protected area effectiveness in Madagascar using three satellite-based datasets – A study using matching methods

Oppiaine /Läroämne – Subject
Metsien ekologia ja käyttö

| Työn laji/Arbetets art – Level<br>Pro gradu -tutkielma | Aika/Datum – Month and year<br>Huhtikuu 2022 | Sivumäärä/ Sidoantal – Number of pages<br>53 |
|---|---|---|


Tiivistelmä/Referat – Abstract

Suojelualueiden tehokkuuden (tehokkuus = suojelualueen arvioitu kyky estää ei-luonnollisista syistä johtuvia häiriöitä) arvioiminen on välttämätöntä, sillä ne ovat tärkeimpiä työkaluja monimuotoisuuskadon hillitsemisessä, ja niiden tehokkuudessa on huomattavaa vaihtelua. Metsäisillä suojelualueilla yleisin muuttuja tehokkuuden estimointiin on metsäkato, mutta myös tulipaloja voidaan käyttää kuvaamaan maankäytön muutosta. Vankkoja vertailuja erilaisten tietoaineistojen ja proxy-muuttujien välillä ei kuitenkaan ole tehty. Tässä tutkielmassa vastataan kyseiseen puutteeseen vertailemalla kolmea satelliittipohjaista aineistoa (metsäpeite/-kato, aktiiviset palot, palanut alue) Madagaskarin suojelualueiden tehokkuuden arvioinnissa. Tutkielmassa vastataan seuraaviin kysymyksiin: tuottavatko kyseiset aineistot ja niistä muodostetut vastemuuttujat yhtäläisiä suojelualueiden tehokkuusestimaatteja ja voisiko aineistoja käyttää ”ristiin” (ovatko ne keskenään vaihdettavissa) tutkimuksessa ja suojelualueiden käytännön hallinnoinnissa. Hypoteesit ovat seuraavat: H1: Aineistot ja niistä muodostetut vastemuuttujat yhtäläisiä tehokkuusestimaatteja, joista kahdella tulimuuttujalla lienee vahvin samankaltaisuus. H2: Aineistoja voi käyttää ristiin niin tutkimuksessa kuin suojelualueiden käytännön hallinnoinnissa.

Vuonna 2005 tai aiemmin perustettujen Madagaskarin suojelualueiden (N=42) tehokkuutta tutkittiin ajanjaksolla 2005–2017. Kolmea binääristä vastemuuttujaa vertailtiin: metsäkato, aktiiviset tulipalot, palanut alue. Lisäksi tarkasteltiin jatkuvaa metsäkatomuuttujaa. Metsäalueita ja koko kaikkia alueita (metsäalueet + muut alueet) tarkasteltiin erikseen. Vastemuuttujien lisäksi analyyseissa käytettiin sopivia ympäristömuuttujia, joilla kontrolloitiin suojelualueisiin kohdistuviin paineisiin vaikuttavia tekijöitä: kaikilla alueilla korkeus, kaltevuus, etäisyys kaupunkeihin, etäisyys vesiväyliin ja vuotuinen sadanta, metsäalueilla näiden lisäksi etäisyys metsänreunaan. Kaikki muuttujat yhdistettiin solukooltaan 1 kilometrin hilaan. *Matching*-menetelmää (lähimmän naapurin Mahalanobiksen etäisyys) käytettiin kontrolliotoksen muodostamiseksi suojelluille soluille (”lohkoille”, *parcels*). Suhteellinen tehokkuus, yhdistetty (*pooled*) suhteellinen tehokkuus ja verkoston (*network*) suhteellinen tehokkuus laskettiin binäärisillä muuttujilla, keskiarvoinen tehokkuus jatkuvalla muuttujalla. Tehokkuusestimaatit tuotettiin koko maalle, biomeille (trooppiset ja subtrooppiset kosteat lehtipuumetsät, trooppiset ja subtrooppiset kuivat lehtipuumetsät, aavikot ja kuivat pensaikot) ja yksittäisille suojelualueille.

Suojelualueet olivat yleisesti vähintään kohtalaisen tehokkaita, ja kaikki muuttujat tuottivat samansuuntaisia, yhtenäisiä tuloksia koko maan ja biomien tasolla, etenkin yhdistetyn suhteellisen tehokkuuden tapauksessa. Keskimäärin kosteiden metsien suojelualueet olivat tehokkaimpia maankäyttöpaineiden torjumisessa, kuivien metsien suojelualueet hieman vähemmän tehokkaita ja kuivien pensaikoiden suojelualueet olivat tehottomimpia. Yksittäisten suojelualueiden tehokkuuksien välillä oli huomattavaa vaihtelua, ja noin puolessa kaikista suojelualueista kaikki muuttujat osoittivat alueen olevan tilastollisesti merkitsevästi ($\alpha = 0{,}05$) tehokas. Hieman yli puolessa suojelualueista osa muuttujista osoitti suojelualueen olevan tehokas, osa taas vaikutuksen olevan päinvastainen eli suojelualue lisäsi maankäytön muutosta. Metsäalueilla yksikään suojelualue ei kuitenkaan lisännyt muutosta kaikilla muuttujilla mitattuna; kaikilla alueilla tällaisia suojelualueita oli yksi. Tehokkuudessa oli pieniä eroja kaikilla tasoilla (koko maa, biomit, yksittäiset suojelualueet) metsäalueiden ja kaikkien alueiden välillä, mutta erot olivat tilastollisesti merkitseviä vain muutamassa tapauksessa.

Tulokset viittaavat siihen, että analysoituja aineistoja voisi käyttää ristiin ainakin maa- ja biomitasolla suojelualueiden tehokkuuden arvioimisessa trooppisilla alueilla. Aineistoja voi käyttää myös yksittäisten suojelualueiden tasolla tietyin varauksin, aineistojen luonteeseen ja käyttäytymiseen perehtyneenä. Aineistot soveltuvat hyvin tutkimukseen: käytännön toimissa metsäpeite/-kato ja aktiiviset palot lienevät käyttökelpoisempia kuin palanut alue, sillä viimeksi mainitussa on muutamia piirteitä, jotka tekevät sen käyttämisestä kahta muuta haastavampaa (se muun muassa vaatii enemmän prosessointia käyttötarkoituksesta riippuen ja sen lataaminen vaatii erillisen ohjelmiston).



Avainsanat – Nyckelord – Keywords
Madagaskar, suojelualueet, suojelualueiden tehokkuus, matching, Mahalanobiksen etäisyys


Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Prof. Markus Holopainen, FT Johanna Eklund


Säilytyspaikka – Förvaringställe – Where deposited
Helsingin Yliopiston Kirjasto – Helda / E-thesis (opinnäytteet) ethesis.helsinki.fi


Muita tietoja – Övriga uppgifter – Additional information

## Table of contents

# 1 Introduction

## *1.1 Background*

Because of human actions, biodiversity is declining at a rate faster than ever before and will continue to do so unless the problem is addressed immediately and drastic actions are taken (IPBES, 2019). Amidst forest loss, habitat conversion, and destruction, protected areas serve as the last fortresses of untouched nature and are regarded as key tools in tackling the ongoing biodiversity loss (Maxwell et al., 2020). There is a push to increase protected area coverage, as the zero-draft of the new CBD aims to protect up to 30 % of the globe (CBD, 2020). However, there are significant differences in the effectiveness (the estimated ability of protected areas to prevent unnatural disturbances) of protected areas in mitigating threats and maintaining populations, depending for example on the funding and management of specific areas. In addition, there is a lack of sufficient information regarding the effectiveness of protected areas even though the importance of providing firm evidence regarding conservation measures has been brought up several times over the last two decades (Pullin & Knight, 2001; Sutherland et al., 2004; Baylis et al., 2016). In order to avoid misallocation of resources, it is essential to study protected area effectiveness using appropriate methods and accurate data.

## *1.2 Theory*

According to IUCN, a protected area is “a clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values” (IUCN Definition 2008). Protected area effectiveness, in turn, can be defined as follows: perfectly effective protected areas “achieve the long-term conservation of nature, at least within their boundaries, ideally beyond by working as an effective network” (Rodrigues & Cazalis, 2020).

Since protected area effectiveness cannot be observed directly, it must be studied by using proxies. Most studies to date have been using forest loss as a proxy (for example Andam et al., 2008; Gaveau et al., 2009; Shah & Baylis, 2015). Forest degradation has also been used alongside forest loss but since it is harder to observe from remote sensing data, there are few studies that include it (but see for example Schleicher et al., 2017). Many studies base the analyses on different remote sensing data sets (for example LandSat imagery) or use pre-processed data sets such as global tree cover and forest loss data by Hansen et al. (2013). There also exist more detailed data for certain countries and regions, Madagascar included (Vieilledent et al., 2018). Hansen’s data is suitable for large-scale studies, but it may be too inaccurate for smaller areas where more nuances are desired.

Different methods have been used to study and quantify protected area effectiveness. Earlier, the effectiveness was studied by simply comparing deforestation inside and outside protected areas or by limiting the assessment inside protected areas (Naughton-Treves et al., 2005). There are several weaknesses in this method. Most importantly, it does not acknowledge the fact that protected areas are rarely randomly established in the landscape. There are several factors that affect their placement, for example the suitability of the area for agriculture, and accessibility, meaning that protected areas often are established in remote areas where also conversion pressures are lower (Joppa & Pfaff, 2009). This means that simple inside-outside comparisons may overestimate the effectiveness of protected areas, not accounting for the effect of the potential biases affecting both the likelihood of a protected area being established and the likelihood of it being converted to alternative land uses (Andam et al., 2008). In addition, the establishment of protected areas sometimes causes deforestation "spillover", meaning that forest loss is just displaced elsewhere (Ewers & Rodrigues, 2008), which is not taken into account in the simple comparisons.

The approaches to evaluating conservation impact that fail to control for observable covariates correlated with both protection and deforestation usually substantially overestimate avoided deforestation (Andam et al., 2008). For this reason, matching methods were introduced to this field of study in the late 2000s. Among the first ones were a study by Andam et al. (2008) and a study by Gaveau et al. (2009). Andam and colleagues measured the effectiveness of protected area networks in reducing tropical deforestation in Costa Rica using nearest neighbour Mahalanobis matching, Gaveau and colleagues studied the same thing in Sumatra using nearest neighbour propensity score matching. Both studies found protected areas to be effective in tackling forest loss, and they also observed that deforestation leakage (spillover) did not take place (Gaveau et al.) or that it was negligible (Andam et al.). Both studies also employed LandSat imagery, additionally Andam and colleagues used aerial photographs since their study period started in the 1950s.

As mentioned earlier, in the field of conservation science, matching studies mostly use forest cover data sets and forest loss as a proxy (for example Joppa & Pfaff, 2011; Nolte et al., 2013; Carranza et al., 2014). The results of these studies are in line with each other: overall, protected areas reduce tropical deforestation; in some cases, there is variation between individual protected areas (Zhao et al., 2019) and protection types (Schleicher et al., 2017). Apart from "direct" forest loss, the ability to avoid forest fires can be used as a proxy for protected area effectiveness (Adeney et al., 2009; Nelson & Chomitz, 2011). However, that is not nearly as common even though forest fires represent deforestation well in the tropics, especially in those areas where slash-and-burn agriculture is common (Kull, 2002; Adeney et al., 2009; Tasker & Arima, 2019; Edwards et al., 2020), and there exist several high-quality, open access fire data sets (for example MODIS fire products provided by NASA). Fire is vastly used as a means to for example clear fields for cultivation, which gives a good indication of land use conversion pressures in a given area (Kull & Lehman, 2020).

However, there are no studies that compare how the different data sets perform in effectiveness evaluations. This is surprising, seeing the global media coverage of forest fires lately, especially about Brazil's Amazon rainforest fires (for example Kelly, 2021), and that the data sets might capture slightly different dimensions of forest loss. This might be especially true in tropical Africa where the main drivers of forest loss are still subsistence farming practices and slash and burn agriculture. Madagascar is especially well suited for such an inspection because it is still common to clear land for small scale agriculture by fire, but previous studies have only focused on changes in the tree cover (Eklund et al. 2016 and 2022).

### *1.3 Study area*

The Republic of Madagascar is an island country in the Indian ocean, located approximately 400 kilometers to the south-east off the African coast. Madagascar is comprised of the main island Madagascar, and several smaller, peripheral islands. Madagascar is the fourth-largest island in the world with its extent of almost 600000 square kilometers. In 2018, there were approximately 26 million inhabitants in the island, the average population density being approximately 47 per square kilometer (World Bank, 2020a). Madagascar's unique nature has developed in isolation for approximately 90 million years (Torsvik et al., 2000) and therefore it has one of the highest levels of endemism in the world as well as a vast variety of different species, both flora and fauna, making it one of the most important biodiversity hotspots in the world (Goodman & Benstead, 2005).

Madagascar's forests are divided into four distinct types: dry, moist, spiny, and mangrove (Vieilledent et al., 2018). Western and central parts of Madagascar are dominated by dry forests, eastern parts by moist forests, southern parts by spiny forests, and West coast by mangroves. The most vulnerable of these four are the dry and moist forests since they are located in the areas with the highest population densities and most suitable for agriculture (WWF, 2020a; WWF, 2020b; Du Puy & Moat, 1996). Partly due to low human population density, spiny forests have a lower habitat destruction and degradation rate when compared to the other forest types (WWF, 2021; Vieilledent et al., 2018; Du Puy & Moat, 1996).

As of 2019, Madagascar had 116 terrestrial protected areas covering approximately 6,5 million hectares (Waeber et al., 2019), which is approximately 11 percent of the total land area. According to Waeber and colleagues, the first conservation sites were established at the beginning of the 20th century, and the number remained relatively low (46) until 1997. Through the 2000s, new policies and programs enabled that number to grow almost threefold (Gardner et al., 2018).

Despite Madagascar having been recognized as a highly important site for nature conservation for a relatively long time, its nature has experienced destruction and degradation so bad that Goodman and Benstead (2005) estimate that only 10 percent of the original natural habitats remain on the island – however, the often repeated claim that originally 90 percent of the island would have been forested is disputed (Scales, 2014). The loss of original habitats has been going on for a long time, and despite conservation efforts it has become faster since the 1950s. Madagascar is one of the world's poorest countries: according to the World Bank, its GDP was 14.084 billion US dollars in 2019 and almost 71 percent of the population was living under the national poverty line in 2012; the COVID-19 pandemic was expected to bring that percentage up to 76,5 % in 2020 (World Bank, 2020b; World Bank, 2020c). Local peoples' livelihoods depend on the over-exploitation of natural resources both for their own use and for international markets (Mongabay, 2014). Over 80 percent of the population live in rural areas and depend on small-scale agriculture for food and charcoal and firewood for energy (Waeber et al., 2019; WWF, 2013).

### *1.4 Aims and research problem*

The aim of this study is to address the lack of performance evaluations of different data products by comparing three satellite-based data sets: forest loss, burned area, and active fire (referred to also as fire incidence). These three data sets may capture slightly different dimensions of the process of forest loss in the tropics: forest loss data shows the land conversion that has been identified as forest loss but does not inform about the cause (logging, burning, etc.), burned area data and active fire data might reflect and capture slightly different stages of land-use change. In addition, the temporal resolution / availability is a key factor: active fire data is available almost immediately and daily, whereas burned area data is a monthly product, and most forest cover change data are updated once a year for public use. From this point of view, active fire data could be most useful especially for practical management purposes. The data sets are compared in terms of their spatial and temporal resolution, the time span they cover, their user friendliness in terms of downloading and pre-processing needs, and ultimately how they affect the results in a matching analysis using Madagascar as a case study.

Research questions:

Q1: Do the three proxies produce similar results regarding PA effectiveness in Madagascar?

Q2: Could these data sets be used (interchangeably) both for science purposes and in practical management of PAs?

Hypotheses

H1: The three proxies produce similar results with the two fire proxies most likely having a stronger relationship.

H2: The data sets can be used interchangeably both for science purposes and in practical management of PAs.

The lack of comparisons between and knowledge about the behaviour of different data sets in effectiveness assessments is the greatest motivation for the chosen subject. I wish that this thesis would bring at least some awareness of this issue, spark more research regarding it, and thus improve the quality of future effectiveness evaluations. As mentioned earlier, protected areas are the main measure in tackling biodiversity loss, and accurate information about their impact is needed in order to correct existing problems (ineffectiveness) within them, avoid wasting conservation resources and allocate them to places where they will be the most beneficial.

# 2 Materials and methods

## *2.1 Data*

The data used in this thesis is divided into four subgroups: response variables, covariates, protected areas (i.e. the "treatment" variable), and other data sets. All the data are open-source and freely available online. The research period is 2005-2017. The motivation for the time period is mainly the availability of the data: MODIS fire incidence data is available from November 11$^{th}$ 2000 to November 30$^{th}$ 2020, MODIS burned area data is available from November 1$^{st}$ 2000 to present, and Vielledent and others' forest cover data is available for 1953, 1973, 1990, 2000, 2005, 2010, 2015, and 2017. Since fire data is not available for the whole year of 2000, the starting point was chosen to be 2005. The ending point, 2017, was chosen because it is the most recent year for which all data sets were available. The analyses were conducted separately for forested areas and the full landscape. Forested areas were defined by all the parcels containing at least one centre point of a cell in the 2005 forest cover layer: such parcels were classified as forest (1), others were classified as non-forest (0).

### *2.1.1 Response variables*

The main data sets analysed in this study are a forest cover data set (Vieilledent et al., 2018), a MODIS active fire data set (Giglio et al., 2016), and a MODIS burned area data set (Giglio et al., 2015). Further details about the response variables are given in Table 1. Regarding the accessibility and usability of the data sets and to provide an opinion about the usability of the data sets in practical management, a user friendliness grade is provided. The grade is subjective and based on my own experience when acquiring the data.

Table 1. Details of the three main data sets.

| Data set | Provider | Access | Available period | Download date | Spatial resolution | Temporal resolution | Information | User friendliness (1=worst, 10=best) |
|---|---|---|---|---|---|---|---|---|
| Forest cover data (Vieilledent et al 2018) | Vieilledent et al 2018 | CIRAD Dataverse, https://dataverse.cirad.fr/dataset.xhtml?persistentId=doi:10.18167/DVN1/AUBRRC | Years 1953, 1973, 1990, 2000, 2005, 2010, 2015 and 2017 | Originally for a former project 23.10.2019; reaccessed 15.12.2020 | 30 m | Yearly | Raster; binary, pixel value 1 indicates forest, 0 non-forest | 10: very easy to find, access, and download. Clear documentation. |
| MODIS C6 burned area data MCD64A1 | NASA | LAADS DAAC official download site (only HDF) or SFTP server of University of Maryland (HDF, GeoTIFF, SHP) | November 2000 to the present (Terra), July 2002 to the present (Aqua) | 12.1.2020 | ~500 m | Daily; distributed as a monthly product | Raster; each pixel contains a day (1-366) of Julian calendar denoting the day that the fire was observed in | 2: relatively easy to find, hard to access and download. Login and FTP software required for downloading. Clear documentation. |
| MODIS C6 active fire data MCD14ML | NASA | NASA archive download tool (for data older than 7 days), https://firms.modaps.eosdis.nasa.gov/download/ | November 2000 to the present (Terra), July 2002 to the present (Aqua) | 11.1.2020 | 1 km | Daily; near-real time product also available | Shapefile; each point denotes a fire and contains information of the brightness of the observation, the date and time, etc. | 8: easy to find, access and download requires giving email address. Clear documentation. |

In the following sections, I present the three datasets as well as covariates in more detail.

**Forest cover**

The forest cover data used in this study was produced by Vieilledent et al. (2018) by combining several earlier forest cover and forest loss data sets that were produced mostly using Landsat images. The original data is in 30-meter resolution. There are also other, global forest cover data sets available (for example Hansen et al., 2013) but they have issues with accuracy and for this study, the forest cover data has to be as accurate as possible as it is not only used as one of the proxies but also as a baseline/reference point for the other response variables. Due to this dual role of the data, a "local", up-to-date option was chosen. The forest cover data was downloaded for the years 2005 and 2017, and the actual variable of interest, forest loss, was calculated using them.

**Fire incidence (MODIS MCD14ML)**

The first fire variable is the MODIS active fire (Giglio et al., 2016; MODIS Collection 6…). It detects fires in 1-km pixels burning at the time of overpass using a contextual algorithm. The data set includes all the observations detected and classified as fire by the MODIS algorithm as points. Each observation has a reliability score ranging from 1 to 100. For research purposes, NASA recommends omitting the low reliability observations (score under 30) (Giglio et al., 2020). Thus, only the nominal and high reliability observations are included in the analyses.

**Burned area (MODIS MCD64A1)**

The second fire variable is the MODIS burned area (Giglio et al., 2015). The MODIS burned area mapping algorithm takes advantage of spectral, temporal, and structural changes such as deposits of charcoal and ash, removal of vegetation, and alteration of the vegetation structure: it locates occurrence of rapid changes in daily surface reflectance time series data and produces the approximate date of burning. The data set includes all the observations detected and classified as burned area by the MODIS algorithm in approximately 500-meter raster cells. The burned cell values represent the day in the Julian calendar in which the burned area observation was made. The cells classified as other than burned area (for example water) and the ones missing information were cleaned from the data.

### *2.1.2 Covariates*

Seven covariates were used: altitude, slope, distance to cities, distance to roads, distance to waterways, and rainfall for both areas, and in addition, distance to forest edge for forested areas. Additionally, biomes and forest were used as exact matching variables. Biomes are as follows:

tropical and subtropical moist broadleaf forests (later moist forests), tropical and subtropical dry broadleaf forests (later dry forests), deserts and xeric shrublands (later shrublands), mangroves, and montane grasslands and shrublands. Further details about the covariates are given in Table 2. The choice of covariates is in line with previous research (for example Eklund et al., 2016).

Table 2. Source data used for creating covariates.

| Covariate | Original data | Original provider | Original data type | Original resolution | Original data access | Original data download date |
|---|---|---|---|---|---|---|
| Altitude | SRTM DEM | NASA/CGIAR-CSI | Raster | 3 arc seconds (~ 90 m) | https://srtm.csi.cgiar.org/srtmdata/ | 25.2.2021 |
| Slope | SRTM DEM | NASA/CGIAR-CSI | Raster | 3 arc seconds (~ 90 m) | https://srtm.csi.cgiar.org/srtmdata/ | 25.2.2021 |
| Distance to cities | Cities | OpenStreetMap | Point | - | https://www.openstreetmap.org/export#map=6/-18.782/46.923 | 5.11.2019 |
| Distance to roads | Roads containing categories "primary", "primary_link", "secondary", "secondary_link", "tertiary", "tertiary_link", "track", "unclassified" | OpenStreetMap | Line | - | https://www.openstreetmap.org/export#map=6/-18.782/46.923 | 5.11.2019 |
| Distance to waterways | Waterways containing categories “river” and “canal” | OpenStreetMap | Line | - | https://www.openstreetmap.org/export#map=6/-18.782/46.923 | 5.11.2019 |
| Distance to forest edge | Forest cover | Vieilledent et al., 2018 | Raster | 30 m | https://dataverse.cirad.fr/dataset.xhtml?persistentId=doi:10.18167/DVN1/AUBRRC | 23.10.2019 |
| Rainfall | IMERG average monthly precipitation (mm) | NASA | Raster | 0,1 degrees (~ 11 km) | https://gpm.nasa.gov/data/directory | 8.3.2021 |
| Biomes | Ecoregions and biomes | Resolve | Polygon | - | https://developers.google.com/earth-engine/datasets/catalog/RESOLVE_ECOREGIONS_2017 | 8.3.2021 |
| Forest | Forest cover | Vieilledent et al., 2018 | Raster | 30 m | https://dataverse.cirad.fr/dataset.xhtml?persistentId=doi:10.18167/DVN1/AUBRRC | 23.10.2019 |

When looking at Table 2, it should be noted that the information is given for the original data sets from which the covariates were then derived with various pre-processing methods which will be discussed in detail in Methods, not for the final covariate layers which were used in the analyses.

The purpose of the covariates is to eliminate the effect of the non-random establishment of protected areas on the effectiveness estimates and make the estimates of PA effectiveness more accurate compared to a conventional approach (simple comparisons between PAs and unprotected areas). If the confounding factors are not taken into account, the effectiveness estimates are usually larger. PAs are usually established in remote, inaccessible, and agriculturally unfit locations, which usually inflates the effectiveness estimates in simple comparisons.

The confounding effect of each individual covariate could be assessed by first running analyses with all covariates, then one by one eliminating those that yield the worst covariate balance, re-running the analyses, and trying different covariate combinations. However, that is rather time-consuming, so in this study, the covariates are chosen based on previous literature. They may not be optimal, but they account for the basic factors which affect PA placement: altitude and slope contribute to the inaccessibility and also suitability for agriculture; distance to cities, roads, waterways, and forest edge mark accessibility; rainfall, especially when paired with slope, works as a proxy for the suitability for agriculture (Ramaharitra, 2012).

### *2.1.3 Protected areas*

Protected area polygons were obtained from the World Database of Protected Areas (UNEP-WCMC and IUCN, 2019). The original layer was downloaded 23.10.2019 and Johanna Eklund cross-checked it with the protected area listings in Goodman et al. (2018) and Domoina Rakotobe at Forum Lafa (the network for terrestrial protected area managers in Madagascar) and kept the protected areas that were mentioned in all sources. The layer contains information about 114 terrestrial protected areas located in Madagascar. Only the protected areas that were established before the study period's start (year 2005 or earlier) were included in the analyses which narrowed the amount down to 42 (Figure 1; Appendix 1). Of those, 25 protected areas belong to the moist broadleaf forest biome, 10 belong to the dry broadleaf forest biome, and 7 belong to the deserts and xeric shrublands biome. Many of the PAs are located in two biomes (in addition to the three biomes mentioned above, there are also mangroves and mountainous grasslands and shrublands) but for simplicity's sake, all PAs are classified according to the listing by Domoina Rakotobe. For example, Andrigitra is located both in montane grasslands and shrublands as well as tropical and subtropical moist broadleaf forests. The latter takes up a larger part of it, and it is classified as belonging to it in the listing.

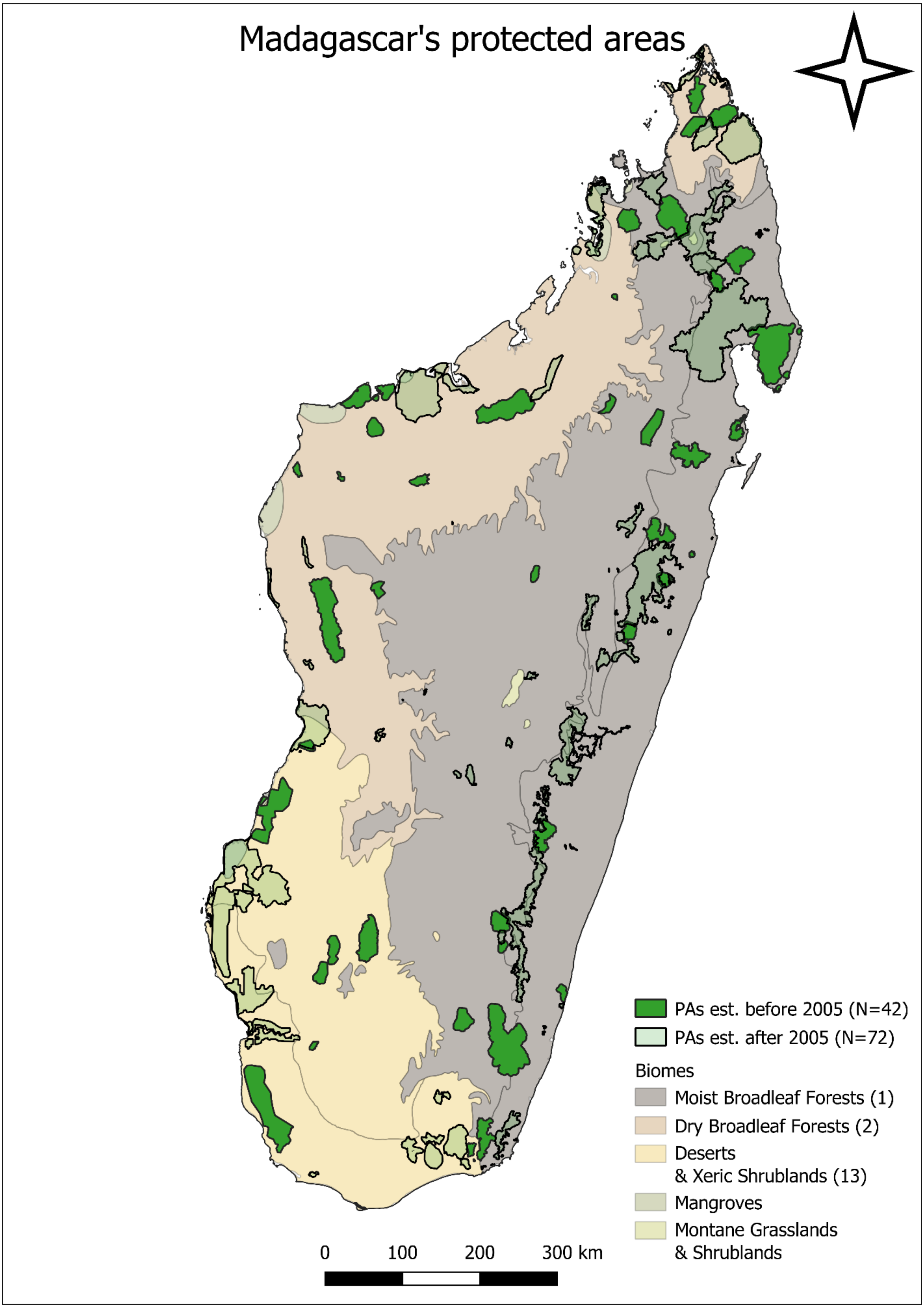



Figure 1. Protected areas of Madagascar. The PAs included in the analyses (those established in or before 2005) are displayed as dark green, the ones omitted (the ones established after 2005) in light green.

### *2.2 Methods*

In this thesis, I am dividing the analysis process into three distinct stages: pre-processing, matching, and effectiveness estimation. Discussing all three in detail is essential to clearly and thoroughly communicate this study to the readers. Common to all variables, they were reprojected into WGS1984/UTM zone 38S before any processing took place.

The analyses were conducted in ArcGIS 10.3.1, QGIS Desktop 3.16.4 with GRASS 7.8.5 (QGIS.org, 2021), R 4.0.2 (R Core Team, 2020), and Excel 2016.

#### *2.2.1 Pre-processing*

The main goal of pre-processing was to create a 1-kilometer grid whose cells (later "parcels") contained information of all the variables listed above. A certain amount of alignment was needed in order to get all the layers into the grid. The amount of alignment is very small for most of the data sets as it is dependent on the spatial resolution of the source data set. It reaches its maximum with burned area, ranging from zero to 247,5 metres (half of the original spatial resolution). In this section, I will go through the steps variable by variable, starting from the response variables.

**Response variables**

All response variables (forest loss, fire incidence, and burned area) are binary: a value of 1 indicates the presence of conversion i.e. forest loss, fire, or burned area, respectively. In addition, a continuous forest loss variable was also created.

*Forest loss, binary and continuous*

Both binary and continuous forest loss are very common proxies in similar studies. At first, only binary forest loss was considered, but since the parcel size (1 km) is relatively large, sole binary variables might produce quite inflated effect estimates especially in smaller PAs. Thus, the continuous forest loss variable was included into the analyses to 1) get an understanding of the actual severity of forest loss and 2) get a reference point towards which to reflect the binary variables. The continuous forest loss was calculated as percentage point change (for example Negret et al., 2020): the difference between proportion of forest cover in 2005 and 2017 inside each parcel. Binary forest loss was derived from the continuous variable by classifying all parcels with the continuous forest loss value < 0 as 1 i.e. loss.

For the percentage point approach, the raster layers of forest cover in 2005 and 2017 were first converted into polygons. The polygon layers were then intersected with the 1-km grid and the information about the forest area in both years inside each parcel was joined to the grid and converted to a percentage. The percentage point change was then calculated as a difference between the forest cover percentage in 2017 and forest cover percentage in 2005 for each parcel (see the formula below):

$$\Delta_{forest} = \left(\frac{A_{2017}}{A_G} \cdot 100\%\right) - \left(\frac{A_{2005}}{A_G} \cdot 100\%\right)$$

where $\Delta_{forest}$ is the percentage point change in forest cover in a parcel, $A_{2017}$ is the forest cover in 2017 in a parcel, $A_{2005}$ is the forest cover in 2005 in a parcel, and $A_G$ is the area of a parcel (1 $km^2$). Negative values indicate forest loss, zero indicates no loss. Positive values would indicate increase in forest cover, but there were no such cells.

*Fire incidence*

Fire incidence data for the whole study period was in one point shapefile that covered all the fire observations from January 2005 to December 2017. All the low-confidence observation points (reliability score below 30) were removed from the data. All parcels containing at least one fire observation were classified as 1 (fire) while others were classified as 0 (no fire).

*Burned area*

The original burned area data is in a raster format in which the cell values represent the date in the Julian calendar in which the observation is made, and the data is provided on a monthly basis. The information was thus in 156 separate layers, from January 2005 to December 2017, and several processing steps had to be taken to convert the data into the desired form. First, Water and Missing Values masks were created from the data: in addition to burned and non-burned cells, MODIS burned area data includes information about cells that the algorithm classifies as water or areas that cannot be classified (missing values). They were selected as separate layers and converted to polygons to be used as masks in the later stages.

After creating the masks, all the layers were reclassified: all values above zero (burned area cells) were classified as 1, others were classified as 0. The layers were then stacked together as a mosaic after which it was once again reclassified (all values above zero as 1, zeros remaining as 0) and the water and missing values were masked away from the layer. Finally, the layer was aggregated into the grid.

**Covariates**

Altitude was resampled into the grid by calculating the mean altitude inside each parcel. A similar approach was used for slope: it was first calculated in 3 arc-second resolution and then resampled.

All the distance layers (distance to cities, distance to roads, distance to waterways) were calculated as simple Euclidean distances. Distance to forest edge was calculated by first inverting the forest cover polygon of 2005 and then calculating the distance from inside the "empty" forest patches to the edges of the inverted polygon.

Rainfall is a monthly data set, so it was first transformed into a yearly basis (cumulative rainfall inside each parcel based on the monthly averages), and then converted to the study period by calculating the mean of all the years 2005-2017.

Biomes were used as they were (polygons covering the whole of Madagascar), and the forest "covariate" used when matching full landscape was defined by the forest cover in 2005: the parcels containing at least one centre point of the original forest cover layer cell were classified as forest (1), others were classified as non-forest (0).

**Protected areas**

Only the protected areas established in the year 2005 or earlier were included in the analyses. The proportion of forest area in each protected area was created by intersecting the protected area polygons with the forest cover polygon of 2005 (this information was later used when interpreting and discussing the results). The parcels intersecting the PAs were classified as 1 (protected), others as 0 (unprotected/control).

### *2.2.2 Matching*

Matching is a counterfactual approach designed to account for the non-random nature of certain phenomena (Rosenbaum & Rubin, 1985), in this thesis' context the establishment of protected areas. There are several matching approaches, one of the most commonly used being nearest neighbour matching with propensity score or Mahalanobis distance as the distance measure. The distance measure means the similarity between treated and non-treated units in terms of their covariate values (Stuart & Rubin, 2011). It can be defined for example as the probability of a given cell to be protected using (binary) logistic regression as in propensity score or as a multi-dimensional generalization of the idea of measuring how many standard deviations away a given point/cell is from the mean of the distribution defined by the given covariates as with Mahalanobis distance.

Applying matching methods can be divided into four steps: first, defining closeness i.e. the distance measure used to determine whether an unit is a good match for another; second, implementing a matching method, given the chosen measure of closeness; third, assessing the quality of the matched samples, and depending on the quality, redoing first two steps until getting well-matched samples; and fourth, analysing the results and estimating the treatment effect (Stuart, 2010).

In this study, matching was done in R using the *MatchIt* package (Ho et al., 2011). In all cases, matching was done without replacement meaning that one control unit could be matched with only one treated unit.

**Nearest neighbour propensity score matching**

The first of the three matching methods that were tested was nearest neighbour propensity score matching (later NN PS matching). Nearest neighbour matching is one of the most common methods, as well as easiest to implement and understand (Rubin, 1973). The simplest way to use nearest neighbour matching is to find each treated unit one control unit with the smallest distance from the treated unit. Here the so-called “greedy” NN matching approach was used, meaning that the matching algorithm was allowed to assign the first suitable control unit it finds to a treated unit, not taking into account its potentially better suitability for some treated unit that might come after the one processed at the moment. This means that the order in which the treated subjects are matched may change the quality of the matches. There is an approach (optimal matching) which avoids the issue by taking into account the overall set of matches when choosing individual matches, but it requires a lot of computational power, and it was deemed sufficient to find well-matched groups which greedy NN matching already does; it is not necessary to find perfectly matched pairs. If that was the case then optimal matching could be implemented (Stuart, 2010).

Propensity score is the probability of a unit being assigned to a particular treatment (protection) given a set of observed covariates. Propensity scores are calculated to all units, both treated and untreated, using logistic regression where the covariates are the predictor variables, and the treatment is the binary predicted variable. Given a binary treatment indicator Z, a response variable r, and background observed covariates X, the propensity score can be defined as a conditional probability as follows:

$$e(x) \stackrel{\text{def}}{=} \Pr(Z = 1 \mid X = x)$$

Propensity scores can also be defined in a more complex way depending on the method that is used to estimate them, but here the common definition is enough.

NN PS matching is a straightforward and rigorous approach, and it has been used in several similar studies (for example Carranza et al., 2014; Schleicher et al., 2017; Negret et al., 2020). There are, however, some disadvantages. In some cases, PS matching has been shown to increase model imbalance, inefficiency, model dependence, and bias (King & Nielsen, 2019). The behaviour of PS matching somewhat depends on the choice of covariates: wrong choices are prone to increase bias and decrease precision (Brookhart et al., 2006). However, in this study's context, PS matching should be an appropriate candidate since the most important confounding factors are most likely included in the analyses (given the previous literature). In addition, assuming that all the main confounding factors are present, the suitability of a given matching method can be assessed by analysing common support: whether there is sufficient overlap in the propensity score distributions of treated and untreated variables. In this study, there mostly was, depending on the biome and whether the analysed area was forest or forest and non-forest. This, also, was a sign that NN PS matching may be implemented.

**Nearest neighbour Mahalanobis distance matching**

Nearest neighbour Mahalanobis distance matching (later NN M matching) uses the same matching approach as was presented above with a different distance measure: Mahalanobis distance. Mahalanobis distance is the distance between a point P and a distribution D, a multi-dimensional generalization of the idea of measuring how many standard deviations away P is from the mean of D. It is defined for a treated unit *i* and a control unit *j* as

$$^{\delta}MD(x_i, x_j) = \sqrt{(x_i - x_j)' S^{-1} (x_i - x_j)}$$

where $x$ is a $p\times 1$ vector containing the value of each of the p included covariates for that unit, $S$ is the covariance matrix among all the covariates, and $S^{-1}$ is the (generalized) inverse of $S$ (Greifer, 2021b).

NN M matching is usually better in creating close pairs than PS matching: when using propensity score, the units may be close in terms of its value but not in terms of the covariates, whereas Mahalanobis distance-paired units will have close values in terms of all of the covariates. In addition, Mahalanobis distance accounts for potential collinearity among the variables considered at each pixel (Eklund et al., 2016). Mahalanobis distance is also commonly used in similar ecological studies (for example Andam et al., 2008). A disadvantage of using Mahalanobis distance is that calipers cannot be directly applied to it. Instead, the calipers must be defined by using propensity score for that and Mahalanobis distance for matching. Here, calipers are not used with Mahalanobis matching as it is not commonplace in similar studies.

### Genetic matching

Genetic matching is not as much a specific matching method as it is a way of specifying a distance measure for another form of matching. Nearest neighbour pair matching is most commonly used. In genetic matching, a genetic search algorithm is used to achieve optimal covariate balance by finding scaling factors for each covariate in a generalized Mahalanobis distance formula (Diamond & Sekhon, 2013).

Genetic matching considers the generalized Mahalanobis distance between a treated unit *i* and a control unit *j* as

$$^{\delta}GMD(x_i, x_j, W) = \sqrt{(x_i - x_j)' \left(S^{-\frac{1}{2}}\right)' W \left(S^{-\frac{1}{2}}\right)(x_i - x_j)}$$

where *x* is a $p \times 1$ vector containing the value of each of the p included covariates for that unit, $S^{-1/2}$ is the Cholesky decomposition of the covariance matrix S of the covariates, and W is a diagonal matrix with scaling factors *w* on the diagonal:

$$W = \begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \dots & \\ & & & w_p \end{bmatrix}$$

Genetic matching is not that common in conservation ecology, but it has been used in some studies (for example Hanauer & Canavire-Bacarreza, 2015). Its main advantage is the optimisation of covariate balance (the degree to which the distribution of covariates is similar across levels of the treatment; Greifer, 2021a). That works usually better with larger data sets. The main disadvantage of genetic matching is that it requires a lot of computational power: the larger the number of units (population size) the evolutionary algorithm uses to solve the optimization problem, the better the overall balance (Sekhon, 2011). Genetic matching was only tested with a couple of individual PAs, and as it demanded vast amounts of time to complete even at the lower limit of the recommended population size of 100 (Greifer, 2021b) and did not cause overall significant improvements in covariate balance when compared to the other two methods, it was ruled out as a potential method.

### Choosing the matching method

Since genetic matching was deemed to be an unfit method due to its time-consuming nature, the two methods that were compared over the whole data set were nearest neighbour propensity score and nearest neighbour Mahalanobis distance matching. NN PS matching was applied with calipers of 0,25 standard deviations (Carranza et al., 2014, Schleicher et al., 2017, Zhao et al., 2019, Negret et al., 2020). To choose the better matching method in terms of covariate balance, the balance statistics were extracted for both methods from the *MatchIt* objects. After this, there are different ways of defining the best method; in this thesis, the following criteria derived from Stuart (2010), Rubin (2001) and Rubin & Thomas (1996) were used:

1. the fewest number of "large" standardized differences of means (greater than 0.25) in each PA
2. the largest amount of good variance ratios (between 0,5-2) in each PA
3. the difference in the absolute values of standardized differences of means per each covariate, then per each PA

The comparisons were done PA-wise and the better method for each PA was decided as a sum of the three criteria. For example, if NN PS matching was better according to criterion 1 but NN M matching was better according to criterion 2 and 3 to a certain PA, NN M matching was given a "point". The assessment procedure was repeated with all the 42 PAs and it clearly indicated that NN M matching is more suitable to the data than NN PS matching. In addition, there were some issues with NN PS matching. To demonstrate the nature of the data, a certain unit could be for example protected, forested, located in moist forests, have an altitude of 520 metres, slope of 6.8 degrees, distance to a nearest city of 76 kilometres, distance to a nearest road of 5.4 kilometres, distance to a nearest waterway of 4.2 kilometres, average yearly rainfall of 2395 millimetres, and a propensity score of 0.005, meaning that the probability of the given cell to be protected is very low according to the NN PS logistic regression. In this case, the regression was wrong about the protection status and the overall classification success rate was only about 50 percent which is very low. Also, NN PS matching does not necessarily produce close pairs on covariates as opposed to NN M matching. These and the criticism that PS matching has received (King & Nielsen, 2019), are the reasons why ultimately, Mahalanobis distance was chosen as the final distance measure for matching.

**Execution**

Matching was done using *MatchIt* package in R. Matching was performed for the individual PAs, and the data were used to construct effectiveness estimates for the three main biomes and the whole country, also. Exact matching was performed for the biomes and variable "Forest" when matching full landscape. To avoid splitting PAs in smaller pieces and thus complicating the result interpretation, all five biomes were used in matching, but the results were calculated only for the three main biomes in accordance to the aforementioned classification. For example, if a PA was located both in moist forests and mangroves and was classified as belonging to moist forests, its parcels in moist forests were matched to control parcels in moist forests, and parcels in mangroves were matched to control parcels in mangroves. However, when calculating the results, all its parcels were analysed according to the PA classification.

To give an example of matching in R, here is a code snippet for semi-automated matching in forested areas for all PAs:

```
MAH_FOR <- lapply(FORPALIST_UP,
                  function(x){
                    matchit(PA_STATUS ~ Altitude + Slope + Distance_cities +
                    Distance_roads + Distance_water + Dist_for_edge + Rainfall,
                    data=x, method="nearest", exact = ~ BIOME_NUMBER,
                    distance="mahalanobis")
                         })
```

`MAH_FOR` is the variable storing the *MatchIt* object. `lapply()` iterates through a list of 42 data frames, each of them containing the information of one PA and all unprotected parcels. Inside `lapply()`, the `matchit()` function executes matching according to the arguments specified in it: first the treatment variable (`PA_status`) is given as the predicted variable; next, predictors i.e. covariates (`Altitude, Slope,` etc.) are defined. The data frame which includes the data to be used in matching is defined with argument `data`: here it is x, denoting each item in the PA data frame list. The rest of the arguments are related to defining the matching method. In the example, nearest neighbour Mahalanobis distance matching is used: `method` shows that nearest neighbour matching is used; `exact` shows that for variable `Biome`, exact matching is used, meaning that each treatment unit has to receive a control that is located in the same biome; `distance` defines the distance measure. If calipers were used, there would be also `caliper` giving its size, and `std.caliper` defining whether the caliper was measured in standard deviations or not.

When executing the `matchit()`-function, pairing, subset selection, and subclassification is performed according to the specifications given in the arguments and a *MatchIt* object is stored into the specified variable, in this case `MAH_FOR`. The object contains all the original information for the treated units and their controls as well as distance (if propensity score is used), weights, and subclass for each unit. In order to calculate the treatment effect, the matched data set has to be extracted from the *MatchIt* object using `match.data()` function. In the analyses, matching and treatment effect calculation was performed separately for forested areas and the full landscape (= forested areas plus non-forested areas).

### *2.2.3 Treatment effect estimation*

For the continuous forest loss variable, a mean effect was calculated. For the binary variables, there were several options on how to calculate the treatment effect: for example, absolute effect, relative effect, and pooled relative effect (for example Carranza et al., 2014; Andam et al., 2008; Gaveau et al., 2009; Joppa & Pfaff, 2011; Nelson & Chomitz, 2011.) Here, relative effect was calculated for individual PAs and a pooled relative effect was calculated per biome and for the whole country. Biome and country-level estimates were also derived by calculating the mean based on the individual PA estimates in the respective areas ("network effect") to compare with the pooled effect. All effectiveness estimates and their confidence intervals were derived by using bootstrapping (Davison & Hinkley, 1997; Canty & Ripley, 2021).

**Description of the effect estimation methods**

Relative effect was calculated as the difference between conversion in a PA and its matched control sample, divided by conversion in the control sample. Here, it is important to understand that the metric represents the relative difference between the number of parcels classified as converted in the protected pool/individual PA and control pool/individual control sample. This means that the smaller the sample is, the more inflated the estimates become. For example, if there is a PA in the data that has one converted parcel and its control sample has two converted parcels, the relative protection effect estimate is -50 %. This is one of the reasons why continuous forest loss was kept in the analyses in addition to the binary variables, working as a "reference point" as it is not sensitive to the number of converted parcels.

A relative metric was chosen since it makes the comparison between units meaningful. Comparing absolute effects would be difficult since the PAs are all different-sized, but by using relative effect, they can be more easily compared. Two different effectiveness estimation methods were used for country- and biome-level estimates: pooled relative effect and network relative effect. Pooled effect was calculated by bootstrapping the estimates from all protected and control parcels on a given level (whole country, biome); all the parcels were treated as if they belonged to two large "pools" of protected and unprotected parcels. Network effect was calculated by bootstrapping mean values from the individual PA effectiveness estimates for each larger level (whole country, biomes). The estimates derived from the two metrics were then compared.

Mean effect was calculated for the continuous forest loss variable by subtracting its absolute values in control parcels from the absolute values in protected parcels and calculating the mean in each area.

For all effects, negative values indicate that protection is effective i.e. there are less converted parcels or less forest loss inside the PA/protected pool than the control samples/control pool. Conversely, positive values indicate that PA/ protected pool has, in fact, experienced more conversion than their matched controls.

### Bootstrapping

Bootstrapping is a resampling method utilising random sampling with replacement. By using bootstrapping, it is possible to assign measures of accuracy (for example bias, confidence intervals) to sample estimates and estimate the sampling distribution of almost any statistic. Since no validation data exist, bootstrapping was used to produce confidence intervals in addition to the estimates by repeatedly sampling the data and calculating the desired statistics from the samples. To get a better understanding of the distribution of the bootstrap samples, see Appendix 2. Each estimate was calculated from 10000 bootstrap samples, and 95 % confidence intervals were calculated for all the estimates by using the 2,5$^{th}$ and 97,5$^{th}$ percentiles from the bootstrap distribution. Bootstrapping was done in R using package *boot* (Davison & Hinkley, 1997; Canty & Ripley, 2021).

### Effect comparisons

The binary variables were compared with each other, and the significance in their differences was determined by examining the overlap of their confidence intervals. The continuous forest loss variable was compared with the binary variables only in terms of the direction of the effect since the comparison of the effect would not have been meaningful.

The comparison between forested areas and the full landscape was done using the fire variables (fire incidence and burned area), and the examinations were conducted "inside the variable", i.e. comparing fire incidence to fire incidence and burned area to burned area. The difference in estimates between forested areas and the full landscape was calculated as "forested areas minus full landscape".

The potential relationship between PA area and effectiveness as well as PA forest percentage and effectiveness was examined by calculating Pearson correlations. The statistical significances of the correlation coefficients were derived by calculating the t-scores for the correlation coefficients and drawing the corresponding two-sided p-values for the t-distribution with n-2 degrees of freedom. The formula for calculating the t-score was the following:

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

where *t* is the t-score, *r* is the correlation coefficient, and *n* is the sample size (42 PAs).

# 3 Results

### *3.1 Whole country and biomes – forested areas*

On country and biome level, the pooled estimates converge with all variables, indicating that protected areas are effective (Figure 2). The only exception are the shrublands, where the confidence interval (CI) of burned area overlaps zero, indicating that the pooled effect of protection in the biome cannot be deemed either positive or negative. Fire variables have the largest range across country and biomes, whereas binary forest loss is more consistent, meaning that the estimates across the levels are further from each other when measured with fire than with forest loss (Figure 2). The strongest pooled effect is -42,0 % in the moist forests measured with fire incidence, weakest -2 % in the shrublands measured with burned area (Figure 2). All variables significantly differ from each other in the whole country and the moist forests, in the dry forests binary forest loss and fire incidence do not differ significantly, and in the shrublands binary forest loss and fire incidence as well as fire incidence and burned area do not differ significantly (Figure 2).

In the whole country and all biomes, all variables produce similar network estimates and do not differ from each other statistically significantly. The network effects are mostly similar to the pooled effects in the whole country and the moist forests, apart from the strikingly large confidence intervals for the network effects (Figure 2; Figure 3). In the dry forests, burned area effect seems to have changed from effective to increase conversion (Figure2; Figure 3). Also, in the shrublands, the network effects differ quite a lot from the pooled effects. The more variation there is between individual PAs in a given area, the more difference there is between the pooled and network effect (Appendix 3, Table A3). However, the differences are not statistically significant at the biome or country level.

Pooled continuous forest loss (Figure 2) indicates much lower effectiveness estimates than binary variables, which is explained by the different effectiveness measure but also the nature of the data (see Methods for more detailed explanation). The strongest effect is -4,4 %-points in the moist forests, weakest -0,7 %-points in the shrublands. Network effect (Figure 3) is the same as pooled effect in the whole country and the moist forests, weaker in the dry forests and stronger in the shrublands. The differences are however not statistically significant.

### *3.2 Individual PAs – forested areas*

At the individual PA level, results are more variable (Figure 4). In 19 PAs out of the 42, all binary variables indicate statistically significantly that the PA in question is effective. In none of the PAs, the effect is significantly reversed with all variables. In 23 PAs, the effects are mixed, meaning that one of the variables indicates effectiveness when the two others indicate increased conversion and vice versa. When comparing only the two fire variables, 22 PAs seem effective, 2 cause increased conversion, and in 19 PAs the effectiveness is mixed. In most cases, fire variables produce a stronger effectiveness estimate than binary forest loss, be it positive or negative. The highest effectiveness is -100 % (found in three PAs, Betampona and Lokobe in the moist forests and Beza Mahafaly in the shrublands), the poorest performance is 200 % (Maningoza, the shrublands), meaning that there are thrice as many parcels that have experienced conversion inside the PA as there are inside the PA's control sample. The cause for these rather inflated estimates is discussed in Methods.

In five PAs, there is no significant difference between the effectiveness estimates derived from any of the binary variables, meaning that all of their CIs overlap. In 9 PAs, only the two fire variables produce significantly similar estimates while the binary forest loss differs from them (Figure 4). In two PAs, binary forest loss produces similar results with both fire variables, but the fire variables differ from each other significantly. In 8 PAs, none of the variables overlap with each other, all indicating different effectiveness levels (Figure 4).

31 PAs are significantly effective measured with continuous forest loss (Figure 5). In three PAs, the effect is reversed. In 8 PAs, the effect is neither significantly positive or negative. The strongest effectiveness is -10,6 %-points in Mananara Nord (moist forests), the strongest conversion increase 2,3 %-points in Tsaratanana (moist forests). Interestingly, in only one PA (Ankarafantsika, dry forests) all binary variables significantly indicate effectiveness but the continuous variable does not (Figure 4; Figure 5).

The effectiveness displayed on map along with the location and size of PAs are found in figures 6 and 7. On first glance, the maps seem to indicate there to be a correlation between PA size and effectiveness: the larger the PA, the more effective it seems to be. However, the correlations are virtually zero between the PA area and effectiveness measured with the different variables, and none of them are statistically significant when $\alpha=0,05$ (Appendix 4). Stronger correlations are found between original forest cover percentage and effectiveness estimates, but those are also moderate at strongest, four of them (forest percentage x fire incidence and continuous forest loss in forested areas, forest percentage x fire incidence and burned area in the full landscape) being statistically significant when $\alpha=0,05$ (Appendix 5).

### *3.3 Full landscape*

Above, only forested areas have been discussed. However, there are several PAs where forest cover percentage is relatively low (in all PAs values are ranging from 2 % to 94 %, 19 PAs with forest cover percentage less than 50 % (Appendix 1)). To take these PAs better into consideration and acknowledge the fact that habitat conversion can take place also outside forests, the analyses were run also by using only the fire variables as response and leaving the distance to forest edge out of covariates. However, as the forested areas are the main focus here, the results for the full landscape are presented in less detail and the figures similar to Figure 2–4 are found in Appendices.

The differences in the pooled estimates derived from fire variables between forested areas and the full landscape are relatively small at the country and biome level (on country level the largest absolute difference is 4,3 %-points for fire incidence, on biome level 8,6 %-points for burned area in the shrublands) (Figure 2; Appendix 6, Figure A6.1). However, the differences are significant in 6 out of 8 cases (burned area in the moist forests and the shrublands). In the network estimates, there are no statistically significant differences between forested areas and the full landscape, and the estimates are very similar (Figure 3; Appendix 6, Figure A6.2).

The variability on PA level is moderate: mean difference for fire incidence is 0,9 %-points and -2,6 for burned area (forested minus all), standard deviation 10,7 and 13,6, range -28,2–29,4 and -54,8–39,1, respectively (Figure 4; Appendix 6, Figure A6.3). Forest percentage and difference both in fire incidence and burned area are moderately correlated (Appendix 7). However, there are only five PAs where the variables significantly differ in direction, meaning that either one or both of them point in different directions in forested areas and the full landscape. In 9 PAs, the fire estimates significantly differ from each other between forested areas and the full landscape, in eight of them the difference being in only one variable and in one of them in both.

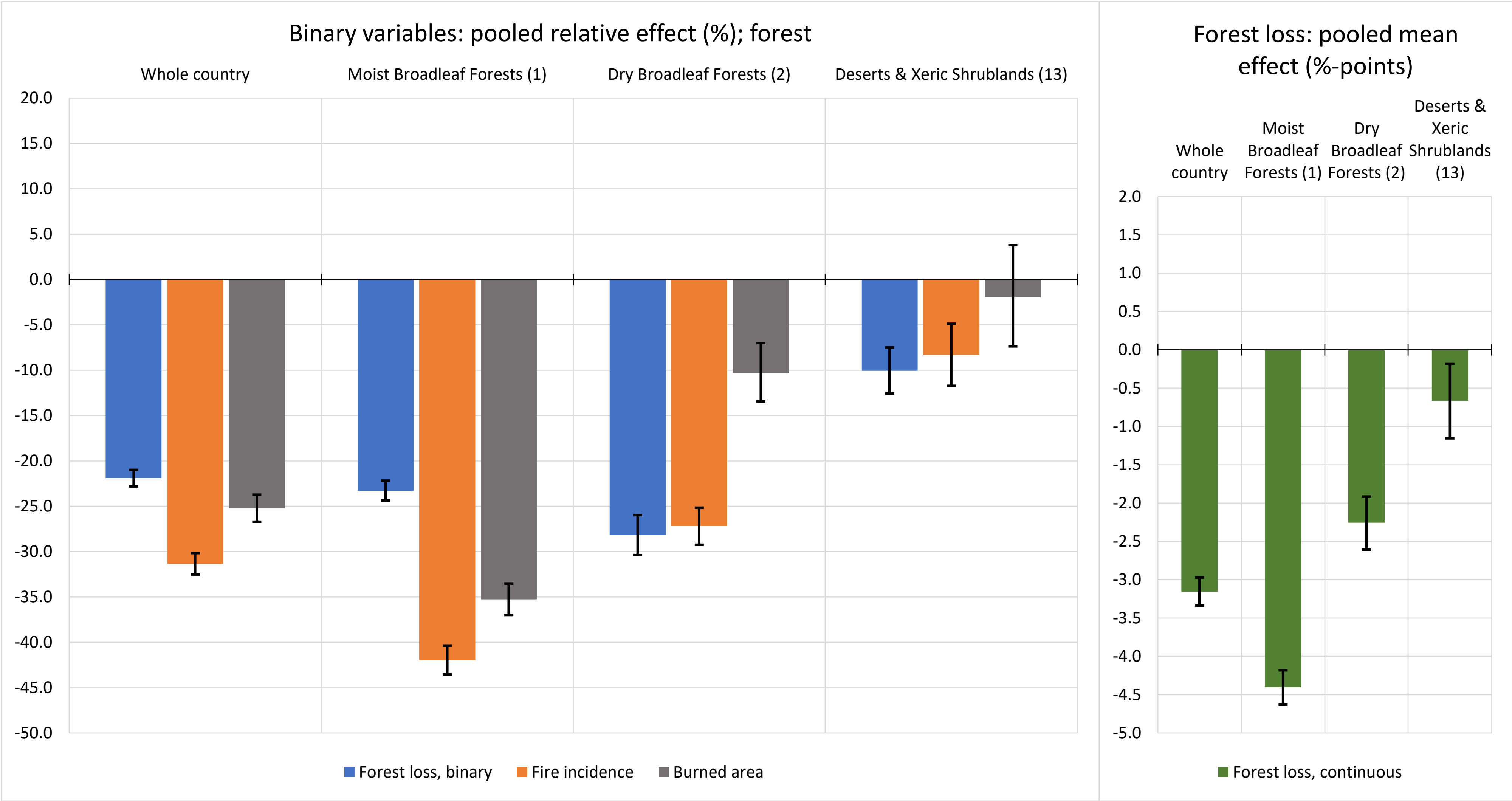



Figure 2. Left: Pooled relative effect (%) on the country and biome level in forested areas, measured with the binary variables. Right: Pooled mean effect (%-points) on the country and biome level in forested areas, measured with the continuous forest loss variable. Note the drastic difference in the scale of effectiveness between binary and continuous variables. The effect measured with binary variables is the relative difference in the amount of parcels labelled as converted on the given level (whole country, individual biomes). The effect measured with continuous variable is the mean difference between relative forest loss in protected parcels and control parcels. In both cases, negative values indicate that conversion is lower in protected areas than control areas, and positive values indicate that conversion is higher in protected areas than control areas. The black bars represent the 95 % confidence intervals of the estimates.

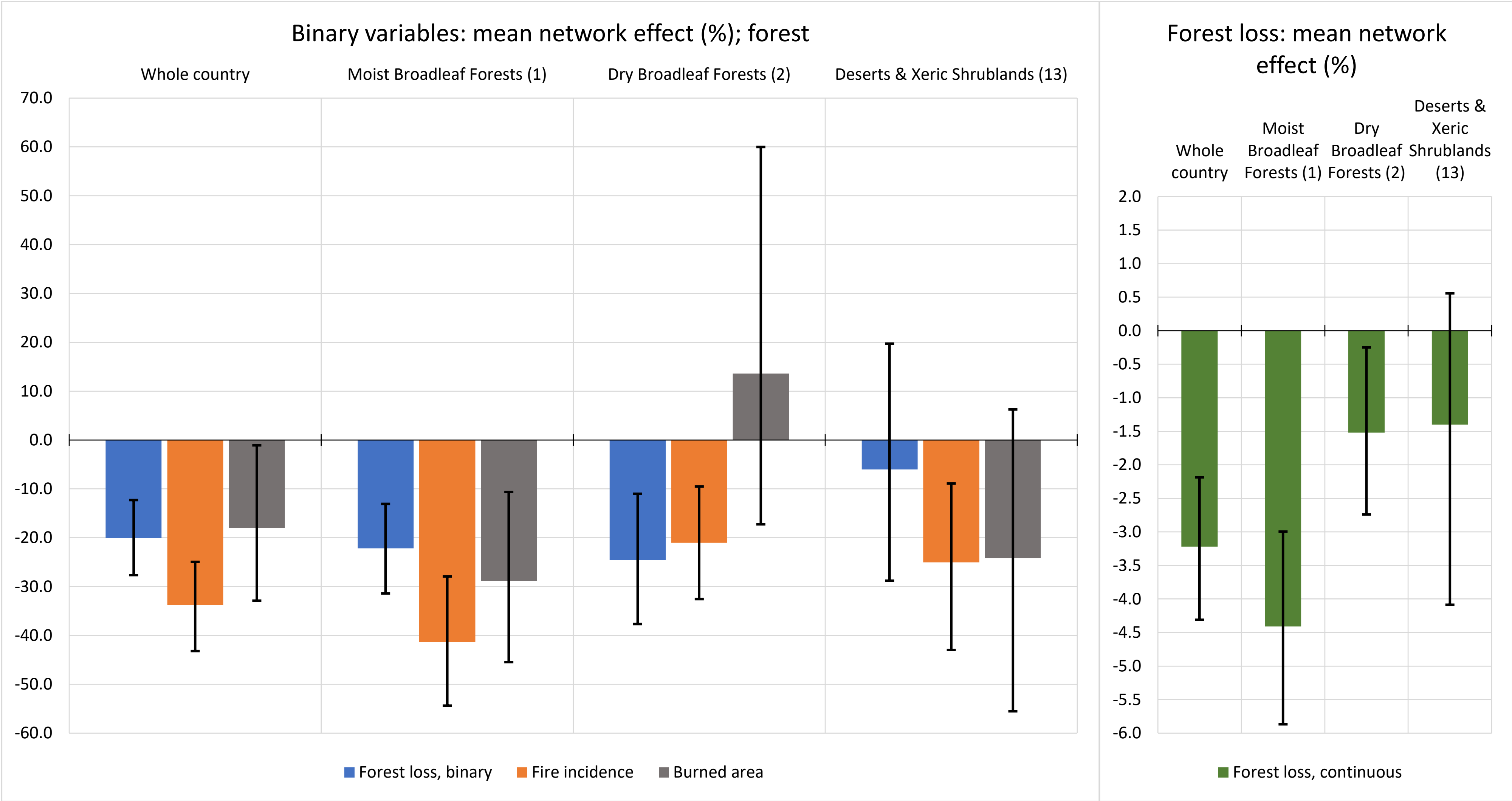



Figure 3 Left: Network effect (%) calculated as a mean of the individual PA effects on the country and biome level in forested areas, measured with the binary variables. Right: Network effect (%-points) on the country and biome level in forested areas, measured with the continuous forest loss variable. Note the drastic difference in the scale of effectiveness between binary and continuous variables. The effect measured with the binary variables as well as the effect measured with the continuous variable is the mean of the individual PA effects on the given level (whole country, individual biomes). In both cases, negative values indicate that conversion is lower in protected areas than control areas, and positive values indicate that conversion is higher in protected areas than control areas. The black bars represent the 95 % confidence intervals of the estimates.

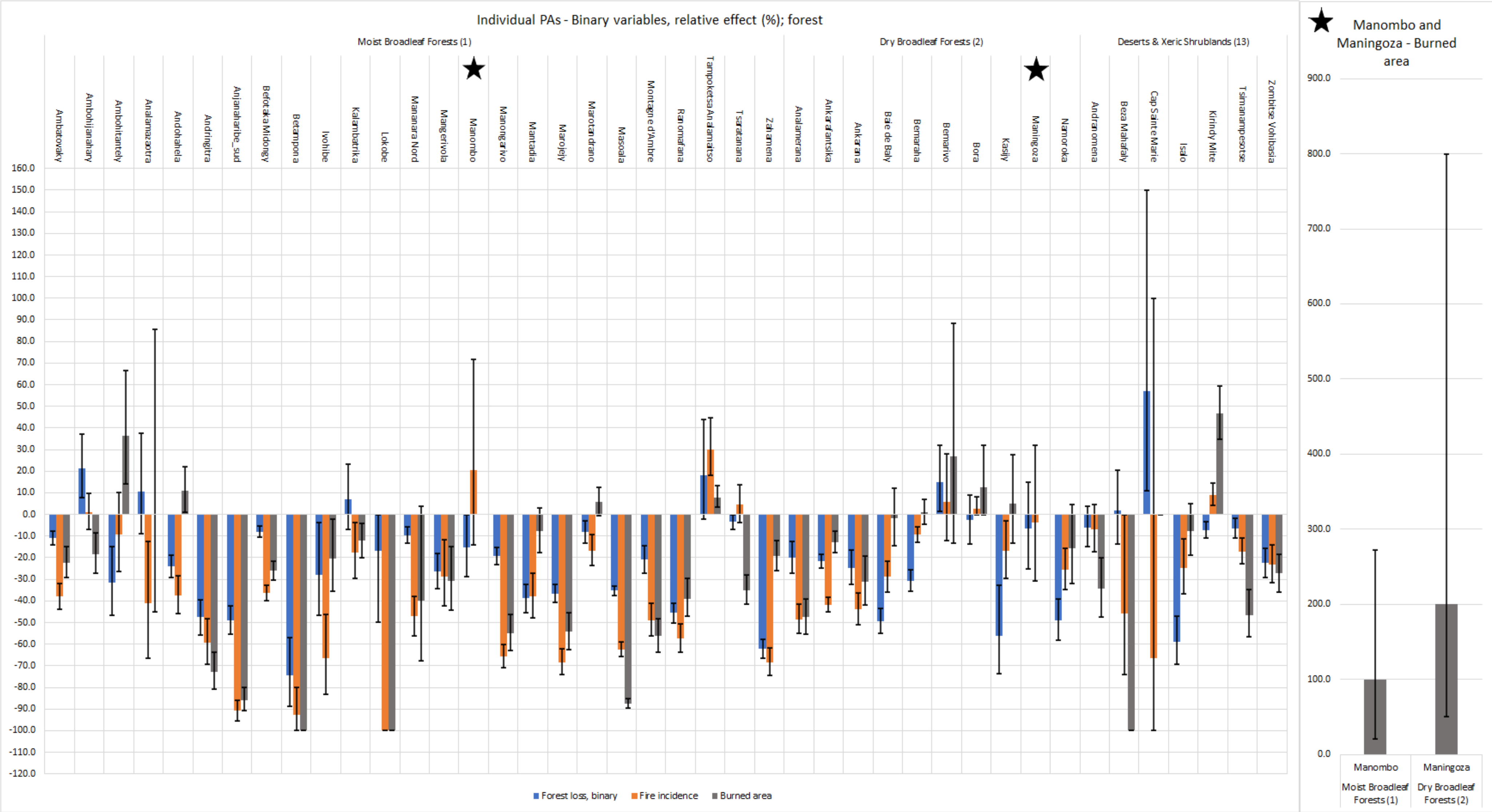



Figure 4. Left: Relative effect (%) on individual PA level in forested areas, measured with the binary variables. The black bars represent the 95 % confidence intervals of the estimates. The black star marks the two PAs where the confidence interval for burned area was very large compared to other PAs and variables, so those estimates are displayed separately. Right: Relative effect (%) measured with burned area in the two PAs where the CI was very large.

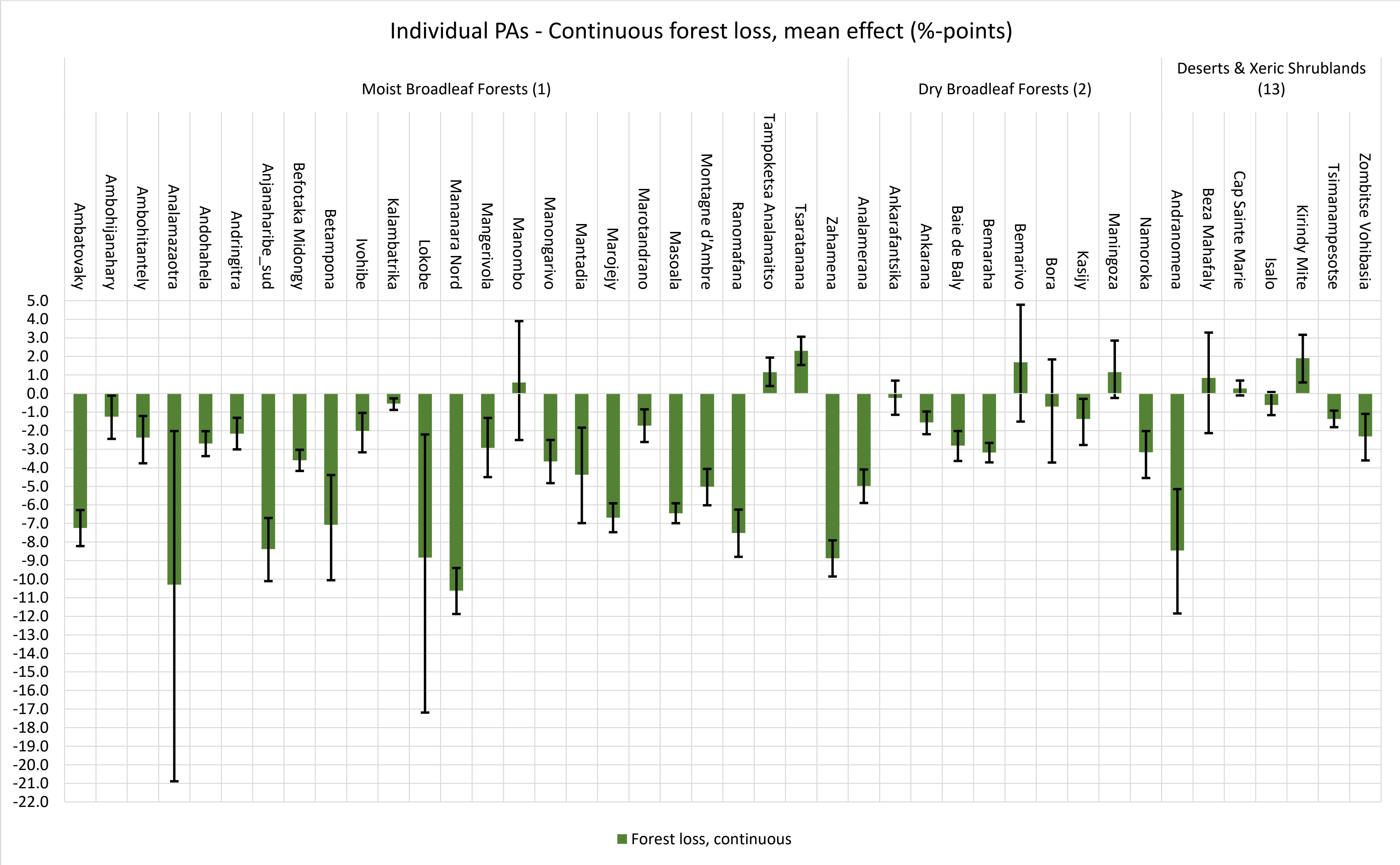



Figure 5. Mean effect (%) of individual PAs measured with continuous forest loss (%-points). The black bars represent the 95 % confidence intervals of the estimates.

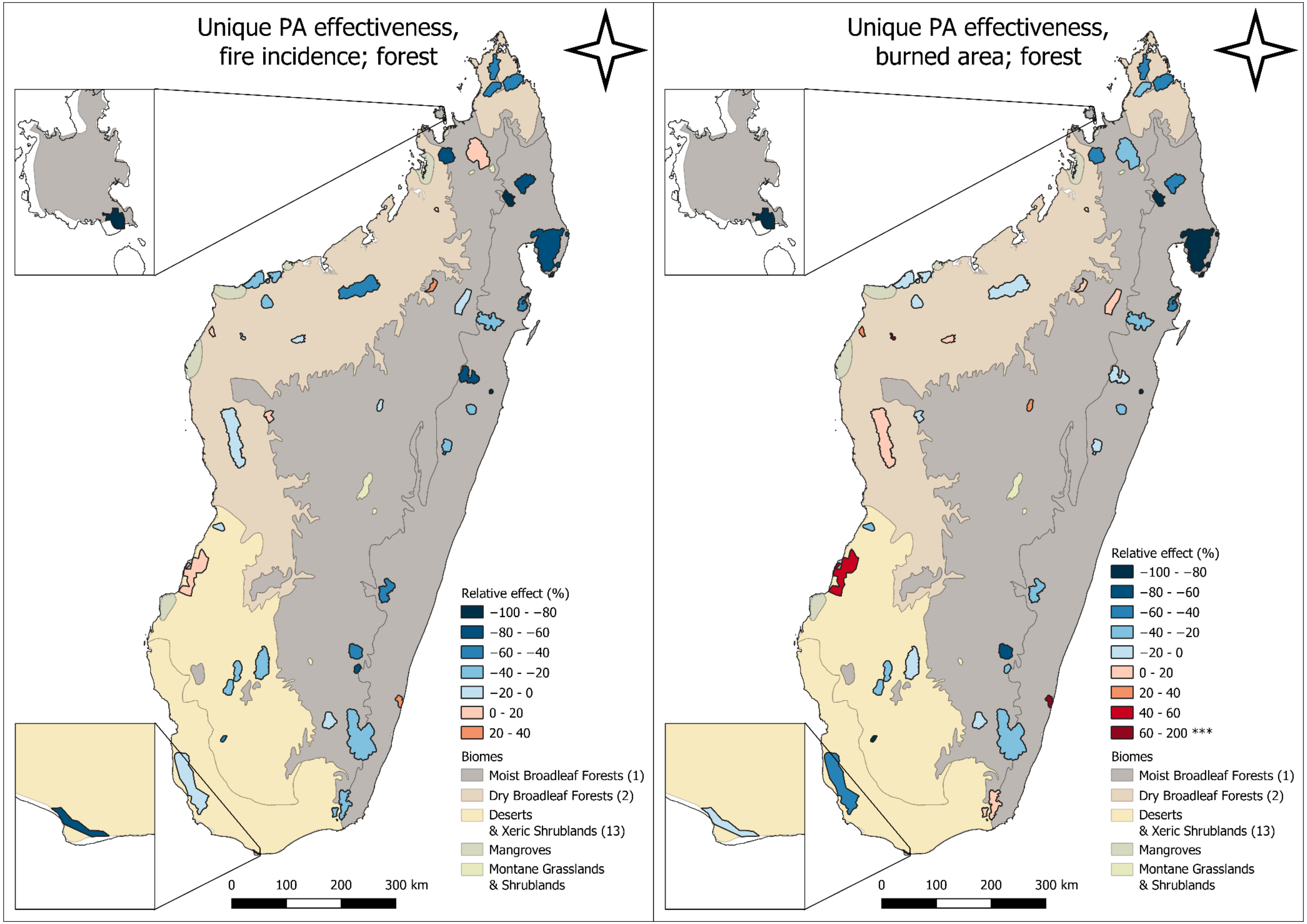



Figure 6. Left: Relative effect measured with fire incidence. Right: Relative effect measured with burned area. *** Note the leap in classification.

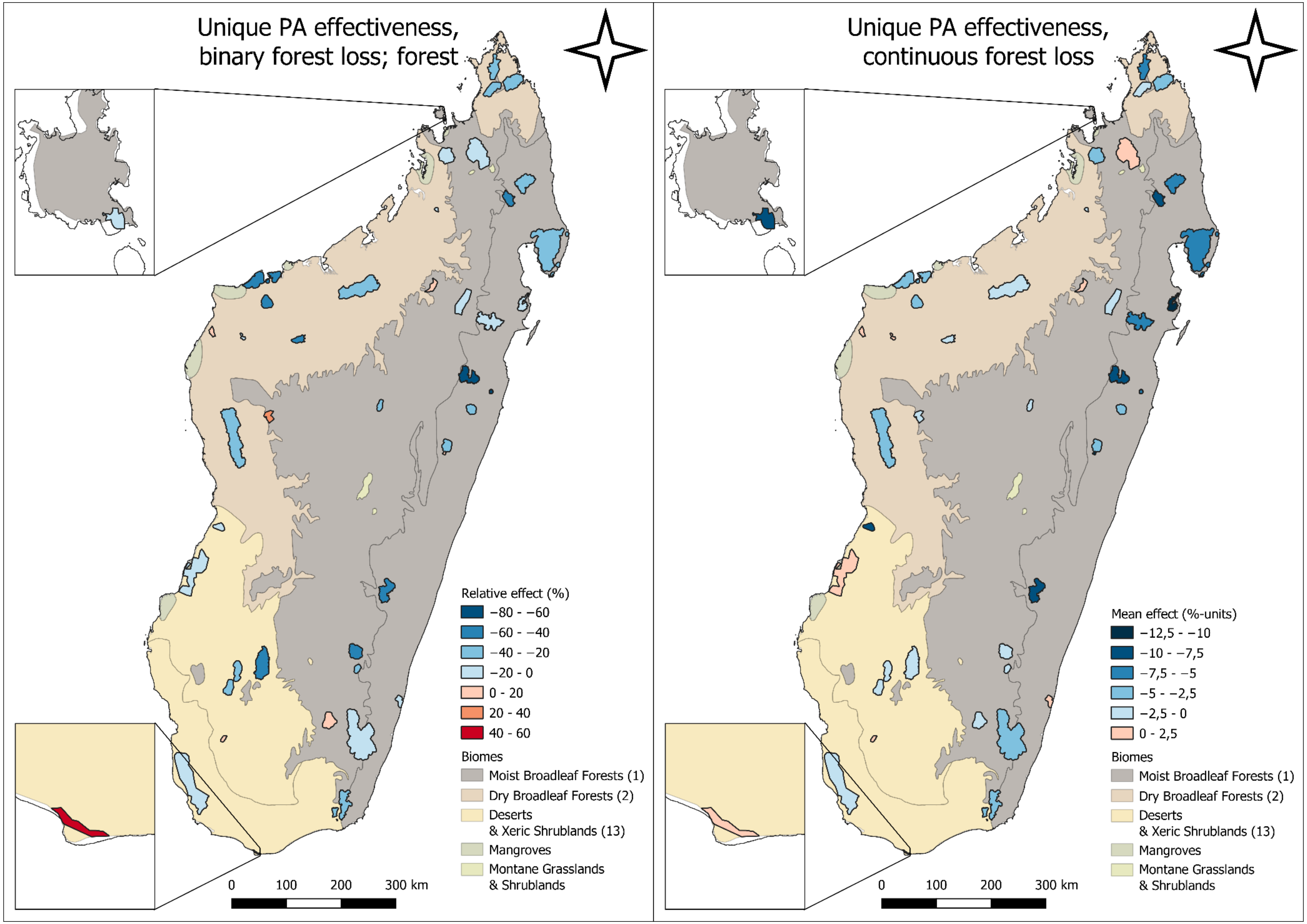



Figure 7. Left: Relative effect measured with binary forest loss. Right: Mean effect measured with continuous forest loss.

## 4 Discussion

On country and biome level, the pooled effect of protection appears to be at least moderately effective regardless of the response variable and area, excluding burned area in the shrublands. The network effect is a bit more variable: burned area produces inconclusive results in the dry forests and shrublands, and binary forest loss in the shrublands. Overall, confidence intervals are very large for the network effects which is mostly caused by the smallish sample size (country level: 42 PAs; moist forests 25 PAs, dry forests 10 PAs, shrublands 7 PAs) and in some cases, the skewness of the bootstrap distribution (see Appendix 2). It seems that at least, if calculating pooled effect on larger areas, any of the variables could be used. In moist forests, PAs are more effective at avoiding fires than forest loss: this could be explained by the fact that deforestation in the rainforests is not only caused by burning but also by logging and harvest of hardwood timber. In the dry forests, forest loss and fire incidence give same effectiveness estimates, as the causes for deforestation are probably more related to burning practices. Interestingly, in the whole country and all biomes, fire incidence produces higher effectiveness estimates than burned area. This could indicate that not all fires lead to burned area i.e. are not particularly severe.

The more drastic differences between variables appear on a smaller scale: there is a substantial amount of variation between and inside individual PAs as shown in results. Regarding the hypotheses, especially the fire variables behaved in some cases rather surprisingly. There are eight PAs where they point to completely different directions (Ambohijanahary, Ambohitantely, Andohahela, Marotandrano, and Tsaratanana in moist forests, Bemaraha, Kasijy, and Maningoza in dry forests). However, when taking the confidence intervals into account, only one PA (Andohahela) displayed effects that are completely opposite. Out of 42 PAs, in 19 all binary variables produced statistically significant results indicating PA effectiveness. It is a promising result in a sense that at the individual PA level, the sample sizes were often very small (the smallest individual PA consisted of seven parcels) and the parcel size large: if the study was done using a smaller parcel size and data sets with higher resolution, I suppose that there would be significantly less variation in the PA level.

An interesting perspective to the results is created when looking closer to the ways of burning that people do in Madagascar. The most prominent threat, especially to forest ecosystems, is the traditional slash-and-burn agriculture, tavy, but it is not the only burning practice. Cattle owners burn non-forested grasslands and shrublands before rains to create new pastures for their animals, but they also create smaller burns ranging from 0,01 to 0,2 hectares in size (Kull & Lehman, 2020). Fires are also ignited to fight uncontrolled fires, ignite regeneration, for destructive or criminal purposes, or as a sign of protest (Kull & Lehman, 2020). The effect of pasture burning is potentially captured in the results of this study, as relative effect of fire incidence and burned area correlate positively with the proportion of non-forest but negatively with the proportion of forest in each PA (meaning that the smaller forest proportion a PA has, the more probably it is deemed ineffective by the two fire variables). Also, in forested areas forest loss and fire variables generally correlate positively with each other, so slash-and-burn agriculture might also be captured by the fire variables, at least to some extent. However, the smaller-scale burns (for example protests, small crop clearings) might not be captured in the data.

Besides studying whether the Vielledent and others' forest cover, MODIS active fire, and MODIS burned area produce similar results in PA effectiveness assessment, the goal was to assess whether these data sets could be used both for scientific study purposes and in practical management of PAs. For study purposes, all the data sets seem suitable if a relatively coarse resolution is enough and the study is carefully designed, considering for example using thresholds for the response variables (such as a certain forest cover percentage when classifying forested parcels). It would be particularly interesting to test different thresholds on each variable and see, how much the effect estimates would be affected by the changes. Regarding PA management, all the data sets received a user friendliness grade (Table 1, page 10). Forest cover and active fire data both are easily accessible, simple to process, and "intuitive" to understand which makes them good candidates for practical management purposes. Burned area is harder to access, process, and understand which could negatively affect its usability. However, the time series -like nature of the data set opens up many possibilities, like examining where and when larger fires are ignited and how they spread.

When comparing forested areas with the full landscape, the differences in the variables between the areas were relatively small, and in most cases insignificant. The lack of significant differences might have roots in the definition of "forested": all parcels intersecting at least one centre point of the forest cover layer cells were classified as forest. Using a threshold of 50 % was contemplated but eventually abandoned as it did not have any backup in the literature and neither did any other threshold; creating an artificial threshold was therefore ruled out. Thus, some parcels classified as forest are forested in less than half of their extent.

There are a few other potential caveats to the study. When creating the binary response variables, any value above zero (forest loss, fire incidence points, burned area cells) in the entire study period were considered as conversion. Thus, a parcel with a value 1 i.e. conversion has not necessarily lost all of its original vegetation, had several fires, or burned entirely: the value simply indicates that some degree of conversion had been experienced.

More precise estimates would have most likely been achieved by using data sets with smaller resolution – albeit open-access, high-resolution data sets suitable for this kind of an assessment are quite scarce if not non-existent, and it is common to use relatively coarse parcel size in similar studies, even as coarse as 3 kilometres (for example Schleicher et al., 2017). In addition, increasing the resolution would have demanded using external services for data processing as executing the analyses would not have been feasible on a laptop or a PC. Other data-related potential sources of uncertainty are combining several data sets of different resolutions and from different sources into one analysis, and the fact that no validation data was used. The latter, however, was not necessarily needed as the main goal was to compare the chosen data sets with each other, not to strictly validate the estimates against other data.

Finally, spatial autocorrelation and spillover effect were not accounted for, and no sensitivity analyses were conducted. However, the spillover effect of protection has, in many cases, been observed to be negligible (for example Andam et al., 2008), and the choice of covariates is in line with previous research, which reduces the need for sensitivity analyses.

## 5 Conclusions

On average, protected areas are moderately effective in Madagascar, regardless of the response variable or data set that is used to create the estimates, or the area in which the estimation is done (forests/full landscape). However, there is some variation between and inside different biomes, and substantial variation on individual PA level. The three data sets analysed in this study could be used interchangeably in the country level and even on the biome level with certain precautions, but at the individual PA level they might produce too variable results, at least when using relatively coarse parcel size or resolution. It might be interesting to repeat the analyses using different data products (for example VIIRS instead of MCD14ML) and a finer resolution. However, all data sets have advantages and disadvantages, and they all can certainly be used in research. However, on operational level particularly burned area could be hard to use due to the nature of the data and the difficulty in access. Fire incidence, on the other hand, is easy to access, understand, and process, and as a daily data set, it could be very beneficial when monitoring PA effectiveness.

## Acknowledgements

I am most grateful for the excellent guidance provided by my two supervisors, Prof. Markus Holopainen and Dr. Johanna Eklund. I also thank The Finnish Biological Society Vanamo (Suomen Biologian Seura Vanamo ry) for their 1500 € grant for working with this thesis. Finally, I thank my family for their support during the working and writing process.

## References

Adeney, J. M., Christensen, N. L., Pimm, S. L. (2009). Reserves protect against deforestation fires in the Amazon. PLoS ONE 4: e5014. https://doi.org/10.1371/journal.pone.0005014

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. Proceedings of the National Academy of Sciences of the United States of America 105: 16089–16094. https://doi.org/10.1073/pnas.0800437105

Baylis, K., Honey-Rosés, J., Börner, J., Corbera, E., Ezzine-de-Blas, D., Ferraro, P. J., Lapeyre, R., Persson, U. M., Pfaff, A., Wunder, S. (2016). Mainstreaming impact evaluation in nature conservation. Conservation Letters 9: 58–64. https://doi.org/10.1111/conl.12180

Brookhart, A., Schneeweiss, S., Rothman, K., Glynn, R., Avorn, J., Sturmer, T. (2006). Variable selection for propensity score models. American Journal of Epidemiology 163: 1149–1156. https://doi.org/10.1093/aje/kwj149

Canty, A., Ripley, B. D. (2021). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.

Carranza, T., Balmford, A., Kapos, V., Manica, A. (2014). Protected area effectiveness in reducing conversion in a rapidly vanishing ecosystem: The Brazilian Cerrado. Conservation Letters 7: 216–223. https://doi.org/10.1111/conl.12049

CBD. (2020). Zero draft of the post-2020 global biodiversity framework. Note by the co-chairs. UN Doc Symbol CBD/WG2020/2/3

Davison, A. C, Hinkley, D. V. (1997). Bootstrap methods and their applications. Cambridge University Press, Cambridge. ISBN 0-521-57391-2. http://statwww.epfl.ch/davison/BMA/.

Diamond, A., Sekhon, J. S. (2013). Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. The Review of Economics and Statistics 95: 932–945. https://doi.org/10.1162/REST_a_00318

Du Puy, D., Moat, J. (1996). A refined classification of the primary vegetation of Madagascar based on the underlying geology: using GIS to map its distribution and to assess its conservation status. Biogéographie de Madagascar. 1996.

Edwards, R., Naylor, R., Higgins, M., Falcon, W. (2020). Causes of Indonesia's forest fires. World Development 127: 104717. https://doi.org/10.1016/j.worlddev.2019.104717

Eklund, J., Blanchet, F. G., Nyman, J., Rocha, R., Virtanen, T., Cabeza, M. (2016). Contrasting spatial and temporal trends of protected area effectiveness in mitigating deforestation in Madagascar. Biological Conservation 203: 290–297. https://doi.org/10.1016/j.biocon.2016.09.033

Eklund, J., Jones, J. P. G., Räsänen, M., Geldmann, J., Jokinen, A-P., Pellegrini, A., Rakotobe, D., Rakotonarivo, O. S., Toivonen, T., Balmford, A. (2022). Elevated fires during COVID-19 lockdown reveal the vulnerability of protected areas. Nature Sustainability (*In press*).

Ewers, R. M., Rodrigues, A. S. L. (2008). Estimates of reserve effectiveness are confounded by leakage. Trends in Ecology & Evolution 23: 113–116. https://doi.org/10.1016/j.tree.2007.11.008

Gardner, C. J., Nicoll, M. E., Birkinshaw, C., Harris, A., Lewis, R. E., Rakotomalala, D., Ratsifandrihamanana, A. N. (2018). The rapid expansion of Madagascar's protected area system. Biological Conservation 220: 29–36. https://doi.org/10.1016/j.biocon.2018.02.011

Gaveau, D. L. A., Epting, J., Lyne, O., Linkie, M., Kumara, I., Kanninen, M., Leader-Williams, N. (2009). Evaluating whether protected areas reduce tropical deforestation in Sumatra. Journal of Biogeography 36: 2165–2175. https://doi.org/10.1111/j.1365-2699.2009.02147.x

Giglio, L., Justice, C., Boschetti, L., Roy, D. (2015). MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 12.1.2021. https://doi.org/10.5067/MODIS/MCD64A1.006

Giglio, L., Schroeder, W., Justice, C.O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. Remote Sensing of Environment 178: 31–41.

Giglio, L., Hall, J. V., Justice, C., Schroeder, W. (2020). Modis Collection 6 active fire product user's guide: Revision C.

Goodman, S. M., Benstead, J. P. (2005). Updated estimates of biotic diversity and endemism for Madagascar. Oryx 39: 73–77. https://doi.org/10.1017/S0030605305000128

Goodman, S. M., Raherilalao, M. J., Wohlhause, S. (Eds.) (2018). The terrestrial protected areas of Madagascar: Their history, description, and biota. The University of Chicago Press for Association Vahatra in Antananarivo.

Greifer, N. (2021a). "Assessing balance" [online source]. Accessed 15.7.2021. https://cran.r-project.org/web/packages/MatchIt/vignettes/assessing-balance.html

Greifer, N. (2021b). “Matching methods” [online source]. Accessed 15.7.2021. https://cran.r-project.org/web/packages/MatchIt/vignettes/matching-methods.html#references

Hanauer, M. M., Canavire-Bacarreza, G. (2015). Implications of heterogeneous impacts of protected areas on deforestation and poverty. Philosophical Transactions. Biological Sciences 370: 20140272. https://doi.org/10.1098/rstb.2014.0272

Hansen, M. C., Potapov, P. V., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R. (2013). High-resolution global maps of 21st-century forest cover change. Science 342: 850–853. https://doi.org/10.1126/science.1244693

Ho, D. E., Imai, K., King, G., Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. Journal of Statistical Software 42: 1–28. https://www.jstatsoft.org/v42/i08/.

IPBES. (2019). The global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Bonn, Germany.

IUCN. (2008). "About - Protected areas" [online source]. Accessed 31.1.2021. https://www.iucn.org/theme/protected-areas/about

Joppa, L. N, Pfaff, A. (2009). High and far: biases in the location of protected areas. PLoS ONE 4: e8273. https://doi.org/10.1371/journal.pone.0008273

Joppa, L. N., Pfaff, A. (2011). Global protected area impacts. Proceedings of the Royal Society B: Biological Sciences 278: 1633–1638. https://doi.org/10.1098/rspb.2010.1713

Kelly, B. (2021, September 4). Amazon fires surge anew in Brazil as cleared forest burns. *Reuters.* https://www.reuters.com/world/americas/amazon-fires-surge-anew-brazil-cleared-forest-burns-2021-09-03/

King, G., Nielsen, R. (2019). Why propensity scores should not be used for matching. Political Analysis 27: 435–454. https://doi.org/10.1017/pan.2019.11

Kull, C. (2002). Madagascar aflame: Landscape burning as peasant protest, resistance, or a resource management tool? Political Geography 21: 927-953. https://doi.org/10.1016/S0962-6298(02)00054-9

Kull, C. A., Lehmann, C. E. R. (2020). Fire in Highland Grasslands: Uses, Ecology, and History [Unpublished manuscript]. In *Goodman, S. M. (Ed.). Natural History of Madagascar (2nd edition)*.

Maxwell, S. L., Cazalis, V., Dudley, N., Hoffmann, M., Rodrigues, A. S. L., Stolton, S., Visconti, P., Woodley, S., Kingston, N., Lewis, E., Maron, M., Strassburg, B. B. N., Wenger, A., Jonas, H. D., Venter, O., Watson, J. E. M. (2020). Area-based conservation in the twenty-first century. Nature 586: 217–227. https://doi.org/10.1038/s41586-020-2773-z

MODIS Collection 6 Hotspot / Active Fire Detections MCD14ML distributed from NASA FIRMS. Available on-line [https://earthdata.nasa.gov/firms]. https://doi.org/10.5067/FIRMS/MODIS/MCD14ML

Mongabay. (2014). "Madagascar" [online source]. Accessed 7.12.2020. https://rainforests.mongabay.com/20madagascar.htm

Naughton-Treves, L., Holland, M. B., Brandon, K. (2005). The role of protected areas in conserving biodiversity and sustaining local livelihoods. Annual Review of Environment and Resources 30: 219-252. https://doi.org/10.1146/annurev.energy.30.050504.164507

Negret, P. J., Marco, M. D., Sonter, L. J., Rhodes, J., Possingham, H. P., Maron, M. (2020). Effects of spatial autocorrelation and sampling design on estimates of protected area effectiveness. Conservation Biology 34: 1452–1462. https://doi.org/10.1111/cobi.13522

Nelson, A., Chomitz, K. M. (2011). Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: A global analysis using matching methods. PLoS ONE 6: e22722. https://doi.org/10.1371/journal.pone.0022722

Nolte, C., Agrawal, A., Silvius, K. M., Britaldo, S.S.F. (2013). Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. Proceedings of the National Academy of Sciences of the United States of America 110: 4956–4961. https://doi.org/10.1073/pnas.1214786110

Pullin, A. S., Knight, T. M. (2001). Effectiveness in conservation practice: Pointers from medicine and public health. Conservation Biology 15: 50–54. https://doi.org/10.1046/j.1523-1739.2001.99499.x

QGIS.org. (2021). QGIS Geographic Information System. QGIS Association. http://www.qgis.org

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ramaharitra, T. T. (2012). Human dimension of conservation planning: The case of Madagascar at national and regional scales. University of California, Berkeley.

Rodrigues, A. S. L., Cazalis, V. (2020). The multifaceted challenge of evaluating protected area effectiveness. Nature Communications 11: 1–4. https://doi.org/10.1038/s41467-020-18989-2

Rosenbaum, P. R., Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. The American Statistician 39: 33–38. https://doi.org/10.1080/00031305.1985.10479383

Rubin, D. B. (1973). The use of matched sampling and regression adjustment to remove bias in observational studies. Biometrics 29: 185–203.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. Health Services & Outcomes Research Methodology 2: 169–188. https://doi.org/10.1023/A:1020363010465

Rubin, D. B., Thomas, N. (1996). Matching using estimated propensity scores, relating theory to practice. Biometrics 52: 249–264.

Scales, I. R. (Ed.) (2014). Conservation and environmental management in Madagascar. 1st edition. Routledge. https://doi.org/10.4324/9780203118313

Schleicher, J., Peres, C. A., Amano, T., Llactayo, W., Leader-Williams, N. (2017). Conservation performance of different conservation governance regimes in the Peruvian Amazon. Scientific Reports 7: 1–10. https://doi.org/10.1038/s41598-017-10736-w

Shah, P., Baylis, K. (2015). Evaluating heterogeneous conservation effects of forest protection in Indonesia. PLoS ONE 10: e0124872. https://doi.org/10.1371/journal.pone.0124872

Sekhon, J. S. (2011). Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. Journal of Statistical Software 42: 1–52. https://doi.org/10.18637/jss.v042.i07

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical Science 25: 1–21. https://doi.org/10.1214/09-STS313

Stuart, E. A., Rubin, D. B. (2011). Best practices in quasi–experimental designs: matching methods for causal inference. (2008). In *Osborne, J. W. Best Practices in Quantitative Methods*. SAGE Publications, Inc. https://doi.org/10.4135/9781412995627.d14

Sutherland, W. J., Pullin, A. S., Dolman, P. M., Knight, T. M. (2004). The need for evidence-based conservation. Trends in Ecology & Evolution 19: 305–308. https://doi.org/10.1016/j.tree.2004.03.018

Tasker, K., Arima, E. (2016). Fire regimes in Amazonia: the relative roles of policy and precipitation. Anthropocene 14: 46-57. https://doi.org/10.1016/j.ancene.2016.06.001

Torsvik, T. H., Tucker, R. D., Ashwal, L. D., Carter, L. M., Jamtveit, B., Vidyadharan, K. T., Venkataramana, P. (2000). Late cretaceous India-Madagascar fit and timing of break-up related magmatism. Terra Nova 12: 220–224. https://doi.org/10.1046/j.1365-3121.2000.00300.x

UNEP-WCMC and IUCN. (2019). Protected planet: the world database on protected areas (WDPA) [online source]. Cambridge, UK: UNEP-WCMC and IUCN. Accessed 23.10.2019 www.protectedplanet.net.

Waeber, P., Rafanoharana, S., Andry Rasamuel, H., Wilmé, L. (2019). Parks and reserves in Madagascar: Managing biodiversity for a sustainable future. In *Bakar, A. N, Suratman, M. N. (Eds.). Protected Areas, National Parks and Sustainable Future.* IntechOpen. https://doi.org/10.5772/intechopen.85348

Vieilledent, G., Grinand, C., Rakotomalala, F. A., Ranaivosoa, R., Rakotoarijaona, J. R., Allnutt, T. F., Achard, F. (2018). Combining global tree cover loss data with historical national forest cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. Biological Conservation 222: 189–197. https://doi.org/10.1016/j.biocon.2018.04.008

World Bank. (2020a). "Country profile: Madagascar" [online source]. Accessed 7.12.2020. https://databank.worldbank.org/views/reports/reportwidget.aspx?Report_Name=CountryProfile&Id=b450fd57&tbar=y&dd=y&inf=n&zm=n&country=MDG

World Bank. (2020b). "Madagascar" [online source]. Accessed 7.12.2020. https://data.worldbank.org/country/madagascar

World Bank. (2020c). Poverty & equity brief, Sub-Saharan Africa: Madagascar. April 2020.

WWF. (2013). Sustainable energy access programme. Access to sustainable energy, a positive environmental tool for socio-economic development.

WWF. (2020a). "Eastern Madagascar" [online source]. Accessed 31.1.2021a. https://www.worldwildlife.org/ecoregions/at0117

WWF. (2020b). "Southern Africa: Northwestern Madagascar" [online source]. Accessed 31.1.2021b. https://www.worldwildlife.org/ecoregions/at0202 17

WWF. (2021). "Madagascar spiny thickets" [online source]. Accessed 31.1.2021c https://www.worldwildlife.org/ecoregions/at1311

Zhao, H., Wu, R., Long, Y., Hu, J., Yang, F., Jin, T., Wang, J., Hu, P., Wu, W., Diao, Y., Guo, Y. (2019). Individual-level performance of nature reserves in forest protection and the effects of management level and establishment age. Biological Conservation 233: 23–30. https://doi.org/10.1016/j.biocon.2019.02.024

# Appendices

**Appendix 1.** Protected areas with information about biome, area, forest-% and year of establishment.

| Name | Biome | Est. (year) | Area (km2) | Forest (%) |
|---|---|---|---|---|
| Ambatovaky | Moist Broadleaf Forests | 1958 | 934 | 92 |
| Ambohijanahary | Moist Broadleaf Forests | 1958 | 243 | 45 |
| Ambohitantely | Moist Broadleaf Forests | 1982 | 172 | 8 |
| Analamazaotra | Moist Broadleaf Forests | 1970 | 27 | 46 |
| Andohahela | Moist Broadleaf Forests | 1939 | 872 | 78 |
| Andringitra | Moist Broadleaf Forests | 1937 | 538 | 40 |
| Anjanaharibe-Sud | Moist Broadleaf Forests | 1958 | 377 | 88 |
| Befotaka Midongy | Moist Broadleaf Forests | 1997 | 3111 | 57 |
| Betampona | Moist Broadleaf Forests | 1927 | 42 | 35 |
| Ivohibe | Moist Broadleaf Forests | 1964 | 148 | 30 |
| Kalambatrika | Moist Broadleaf Forests | 1959 | 524 | 35 |
| Lokobe | Moist Broadleaf Forests | 1927 | 8 | 54 |
| Mananara Nord | Moist Broadleaf Forests | 1989 | 302 | 81 |
| Mangerivola | Moist Broadleaf Forests | 1958 | 271 | 78 |
| Manombo | Moist Broadleaf Forests | 1962 | 186 | 19 |
| Manongarivo | Moist Broadleaf Forests | 1956 | 646 | 79 |
| Mantadia | Moist Broadleaf Forests | 1989 | 331 | 70 |
| Marojejy | Moist Broadleaf Forests | 1952 | 757 | 72 |
| Marotandrano | Moist Broadleaf Forests | 1956 | 674 | 55 |
| Masoala | Moist Broadleaf Forests | 1927 | 2938 | 88 |
| Montagne d'Ambre | Moist Broadleaf Forests | 1958 | 590 | 53 |
| Ranomafana | Moist Broadleaf Forests | 1991 | 729 | 59 |
| Tampoketsa Analamaitso | Moist Broadleaf Forests | 1958 | 226 | 13 |
| Tsaratanana | Moist Broadleaf Forests | 1927 | 1497 | 83 |
| Zahamena | Moist Broadleaf Forests | 1927 | 695 | 94 |
| Analamerana | Dry Broadleaf Forests | 1956 | 754 | 33 |
| Ankarafantsika | Dry Broadleaf Forests | 1927 | 1697 | 44 |
| Ankarana | Dry Broadleaf Forests | 1956 | 487 | 22 |
| Baie de Baly | Dry Broadleaf Forests | 1997 | 1013 | 28 |
| Bemaraha | Dry Broadleaf Forests | 1927 | 2315 | 43 |
| Bemarivo | Dry Broadleaf Forests | 1956 | 120 | 60 |
| Bora | Dry Broadleaf Forests | 1956 | 41 | 32 |
| Kasijy | Dry Broadleaf Forests | 1956 | 230 | 25 |
| Maningoza | Dry Broadleaf Forests | 1956 | 60 | 15 |
| Namoroka | Dry Broadleaf Forests | 1927 | 379 | 22 |
| Andranomena | Deserts & Xeric Shrublands | 1958 | 207 | 64 |
| Beza Mahafaly | Deserts & Xeric Shrublands | 1986 | 78 | 55 |
| Cap Sainte Marie | Deserts & Xeric Shrublands | 1962 | 28 | 76 |
| Isalo | Deserts & Xeric Shrublands | 1962 | 1134 | 2 |
| Kirindy Mite | Deserts & Xeric Shrublands | 1997 | 1745 | 71 |
| Tsimanampesotse | Deserts & Xeric Shrublands | 1927 | 2631 | 82 |
| Zombitse Vohibasia | Deserts & Xeric Shrublands | 1997 | 807 | 51 |

**Appendix 2**. Example distributions of the bootstrap samples.

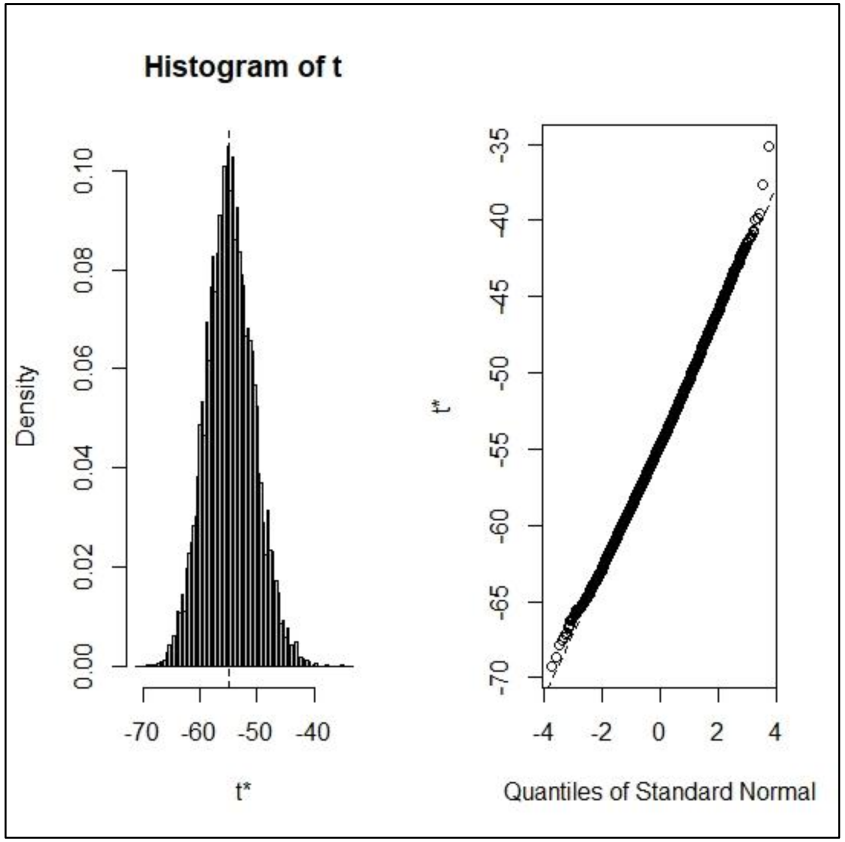



Figure A2.1. Bootstrap sample distribution in Manongarivo (burned area, forested areas). The distribution is normal which was true in most cases. The effectiveness estimate is the expected value of the distribution. *t* represents the estimates drawn from each bootstrap sample (N=10000).

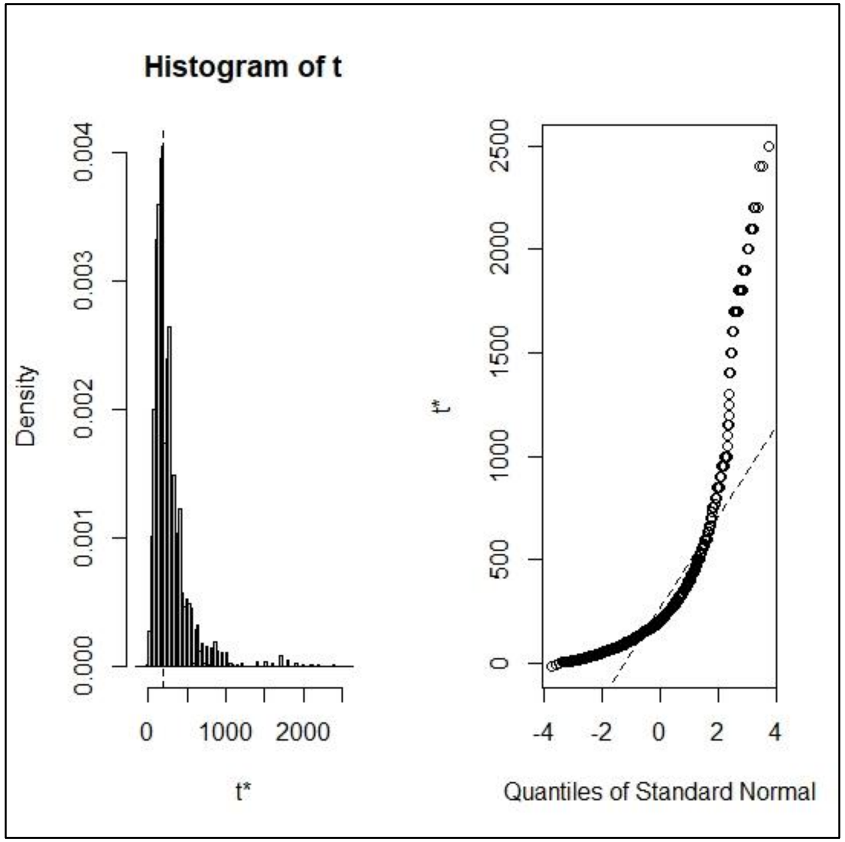



Figure A2.2. Bootstrap sample distribution in Maningoza (burned area, forested areas). The distribution is heavily skewed: there were a couple of such cases in the data, particularly at the individual PA level.

**Appendix 3**. The standard deviations of different variables in the groups (country, biomes) of individual PAs regarding the network effect.

Table A3. The standard deviations of different variables in the groups (country, biomes) of individual PAs regarding the network effect.

| Standard deviations in different PA groups | | | | | | |
|---|---|---|---|---|---|---|
| | | Forest loss, binary | Fire incidence | Burned area | | Forest loss, continuous |
| Forested areas | Country | 25.4 | 30.2 | 52.6 | | 3.5 |
| | Moist forests | 23.3 | 33.2 | 43.8 | | 3.6 |
| | Dry forests | 21.5 | 18.7 | 65.3 | | 2.0 |
| | Shrublands | 31.9 | 23.1 | 41.8 | | 3.2 |
| | | | | | | |
| Full landscape | Country | | 29.4 | 55.9 | | |
| | Moist forests | | 31.6 | 50.1 | | |
| | Dry forests | | 15.1 | 64.6 | | |
| | Shrublands | | 19.1 | 37.6 | | |

**Appendix 4.** Pearson correlations between PA area and effectiveness.

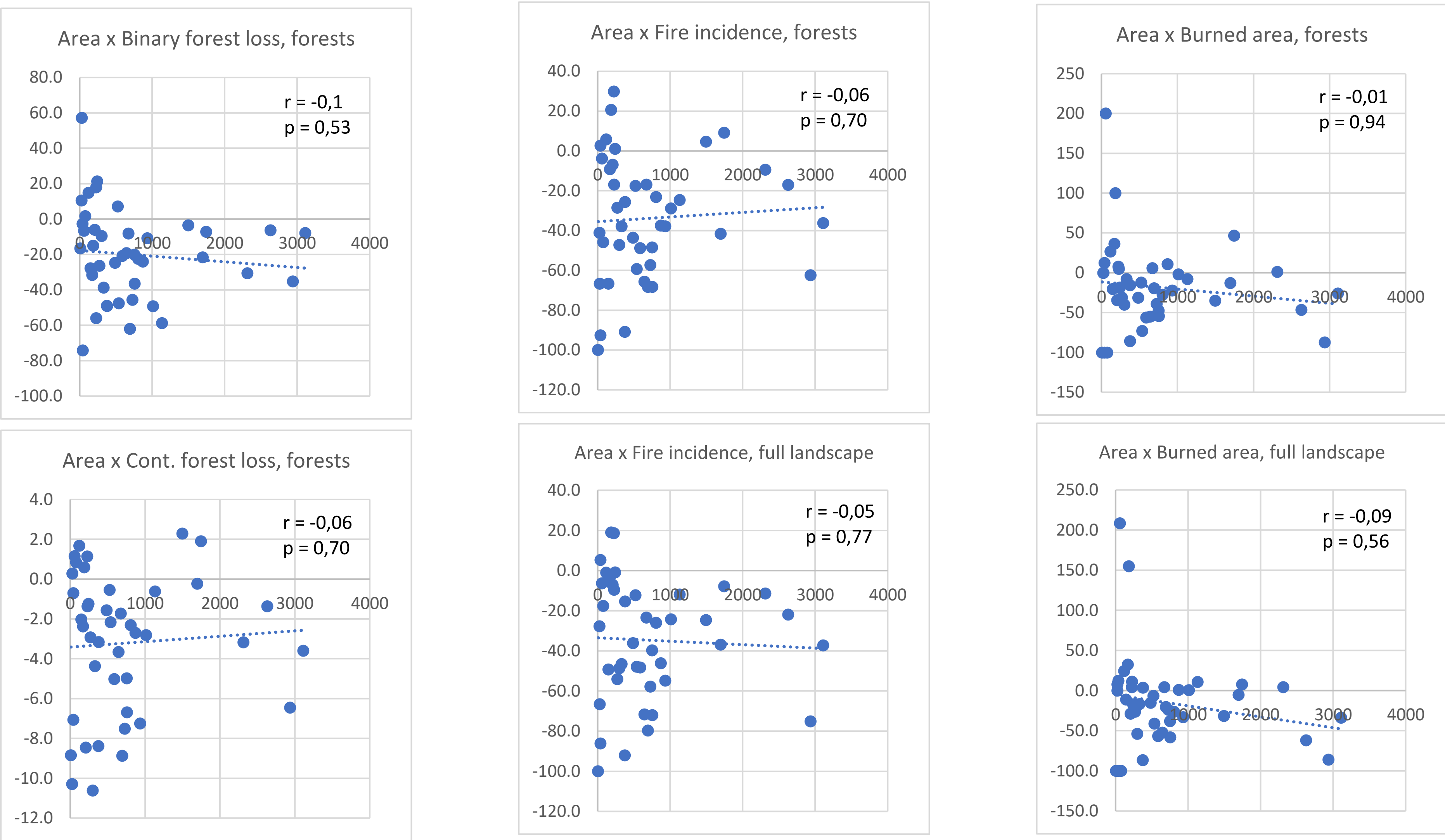



Figure A4. Top row, left to right: 1. area x binary forest loss, forested areas. 2. area x fire incidence, forested areas. 3. area x burned area, forested areas. Bottom row, left to right: 4. area x continuous forest loss, forested areas. 5. area x fire incidence, full landscape. 6. area x burned area, full landscape. In all figures, PA area is on the horizontal axel and the effectiveness is on the vertical axel.

**Appendix 5.** Pearson correlations between PA forest percentage and effectiveness.

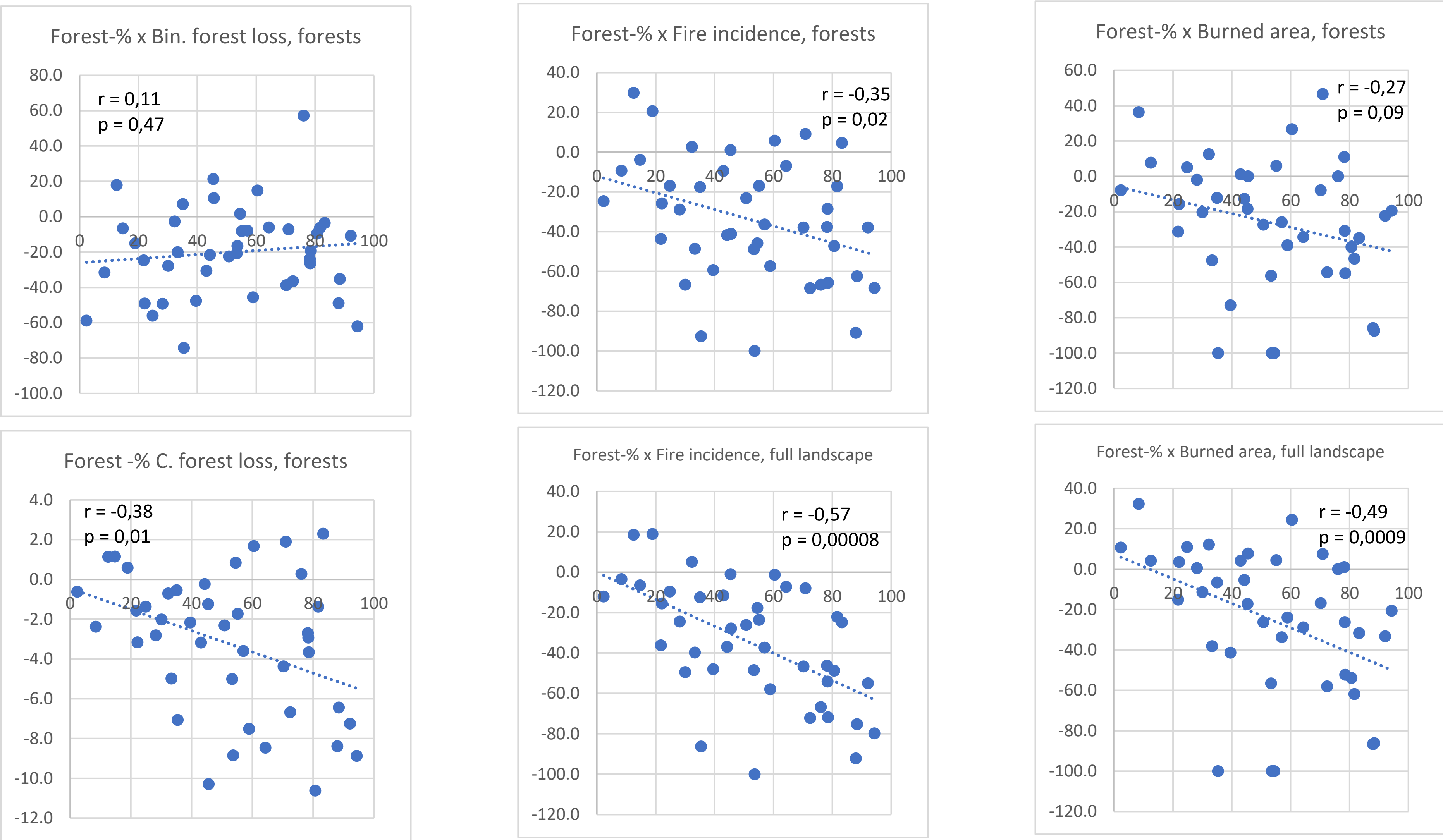



Figure A5. Top row, left to right: 1. forest-% x binary forest loss, forested areas. 2. forest-% x fire incidence, forested areas. 3. forest-% x burned area, forested areas. Bottom row, left to right: 4. forest-% x continuous forest loss, forested areas. 5. forest-% x fire incidence, all areas. 6. forest-% x burned area, all areas. In all figures, PA forest percentage is on the horizontal axel and the effectiveness is on the vertical axel.

**Appendix 6.** Effectiveness figures for the full landscape: whole country and biomes, individual PAs.

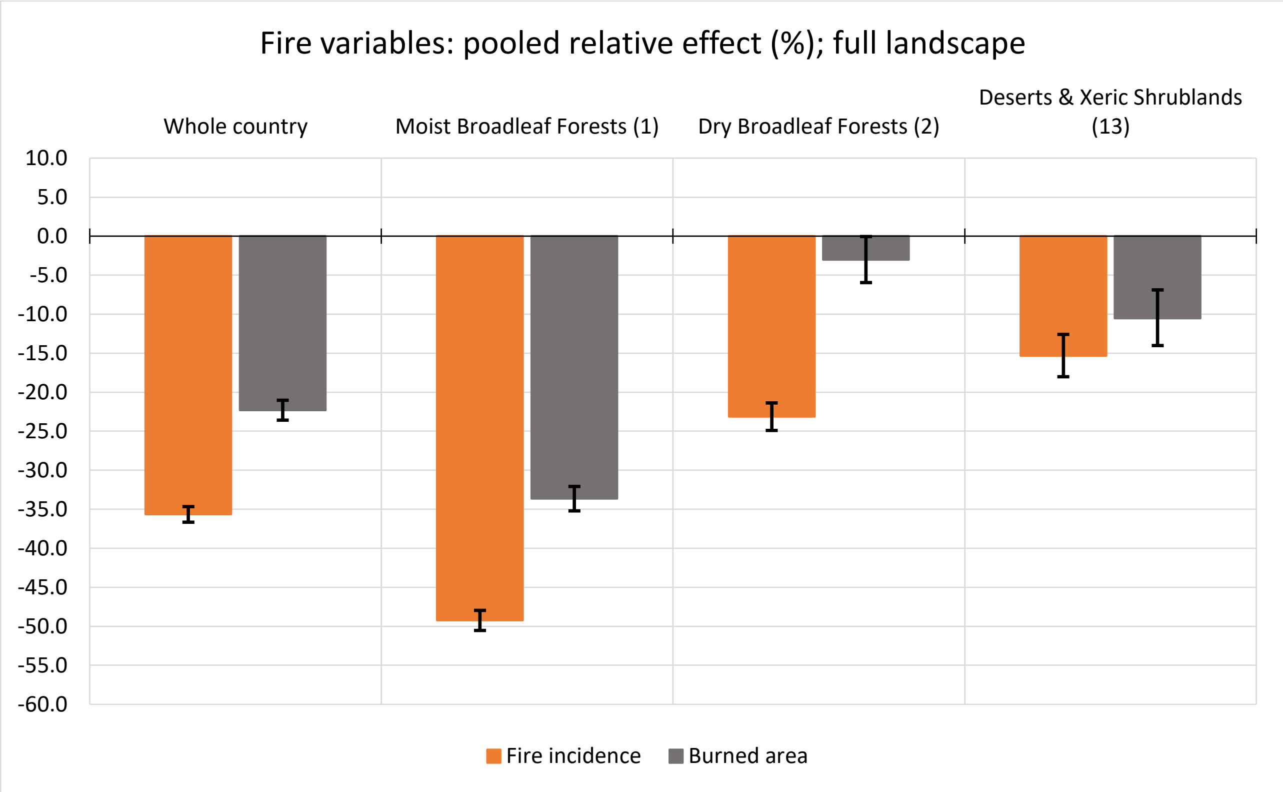



Figure A6.1. Pooled relative effect (%) on the country and biome level in the full landscape, measured with the binary fire variables. The effect is the relative difference in the number of parcels labelled as converted on the given level (whole country, individual biomes). Negative values indicate that conversion is lower in protected areas than control areas, and positive values indicate that conversion is higher in protected areas than control areas. The black bars represent the 95 % confidence intervals of the estimates.

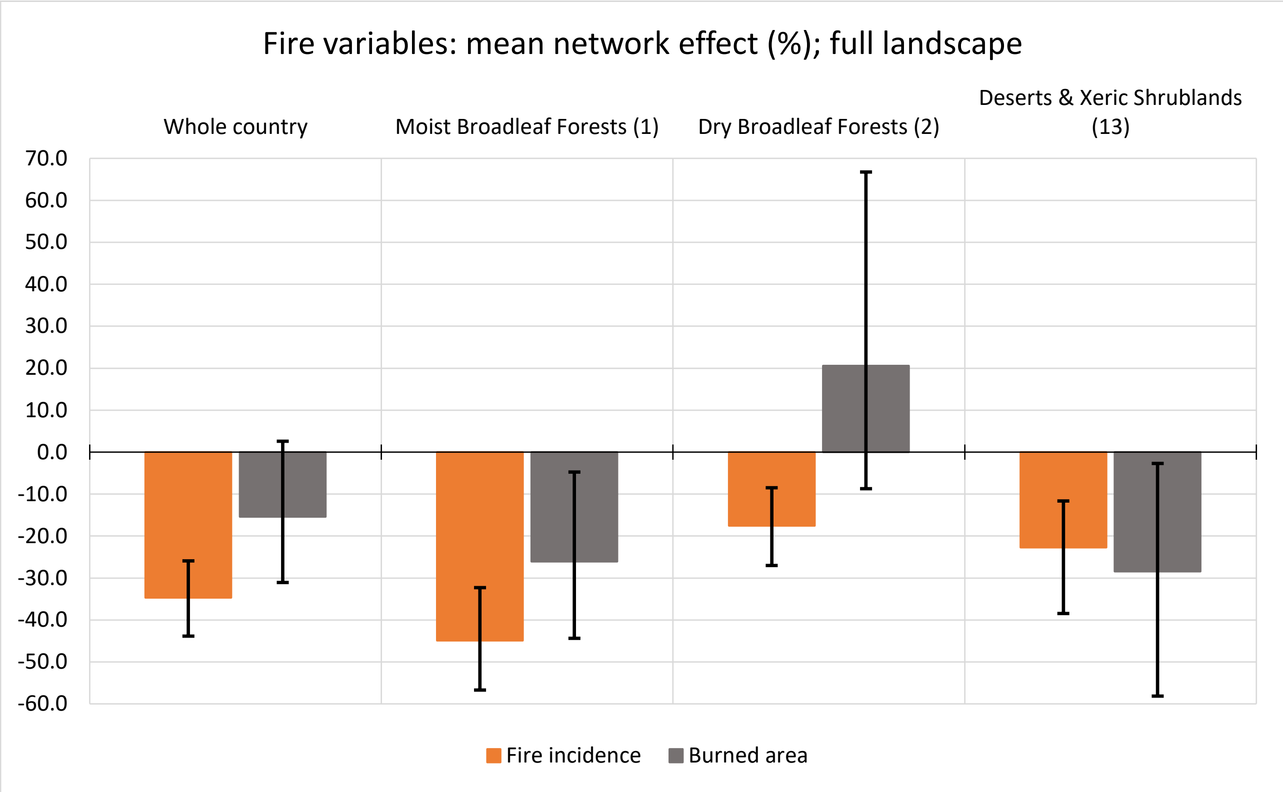



Figure A6.2. Network effect (%) calculated as a mean of the individual PA effects on the country and biome level in the full landscape, measured with the binary fire variables. The effect measured with the binary variables is the mean of the individual PA effects on the given level (whole country, individual biomes). Negative values indicate that conversion is lower in protected areas than control areas, and positive values indicate that conversion is higher in protected areas than control areas. The black bars represent the 95 % confidence intervals of the estimates.

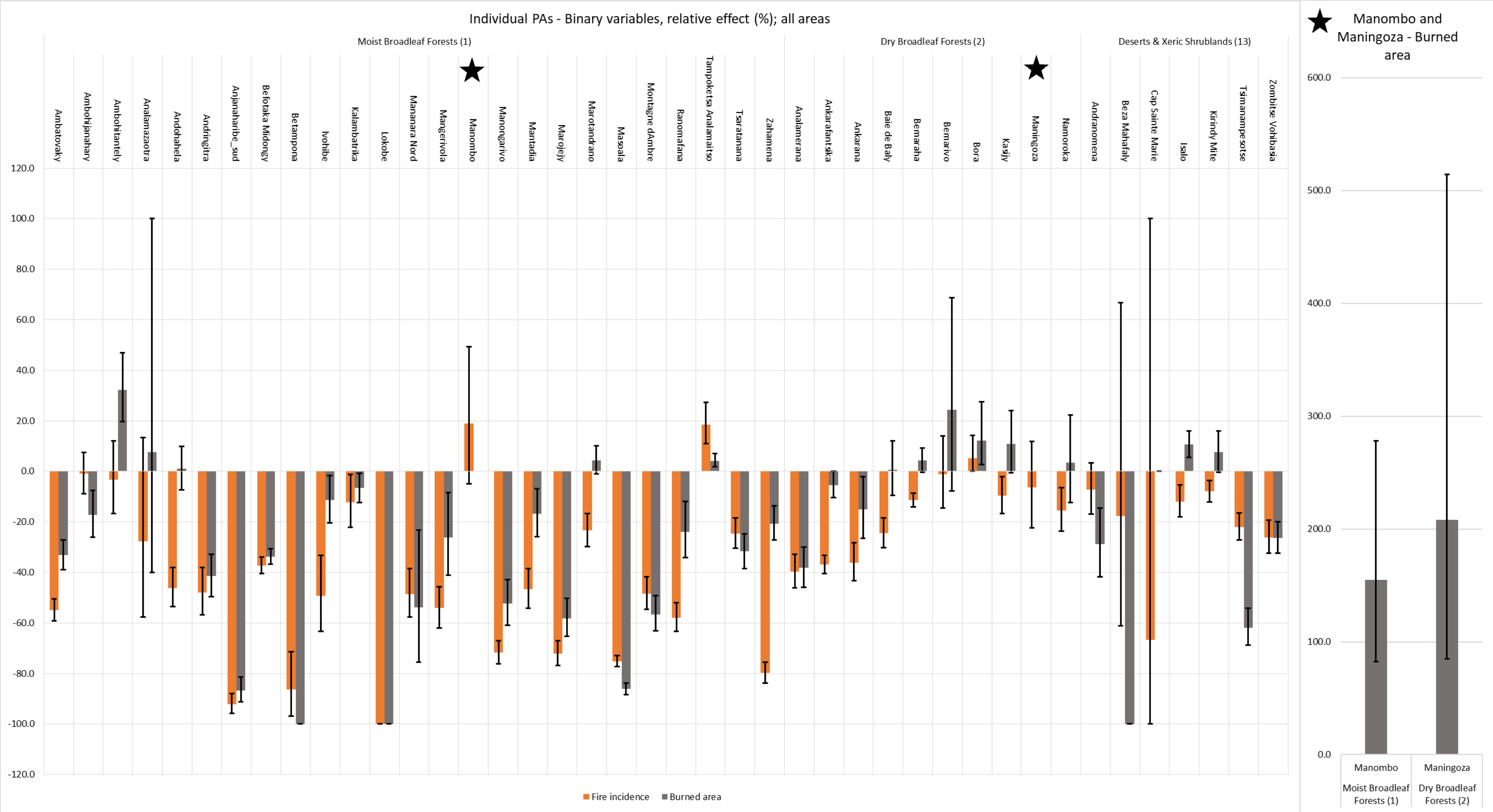
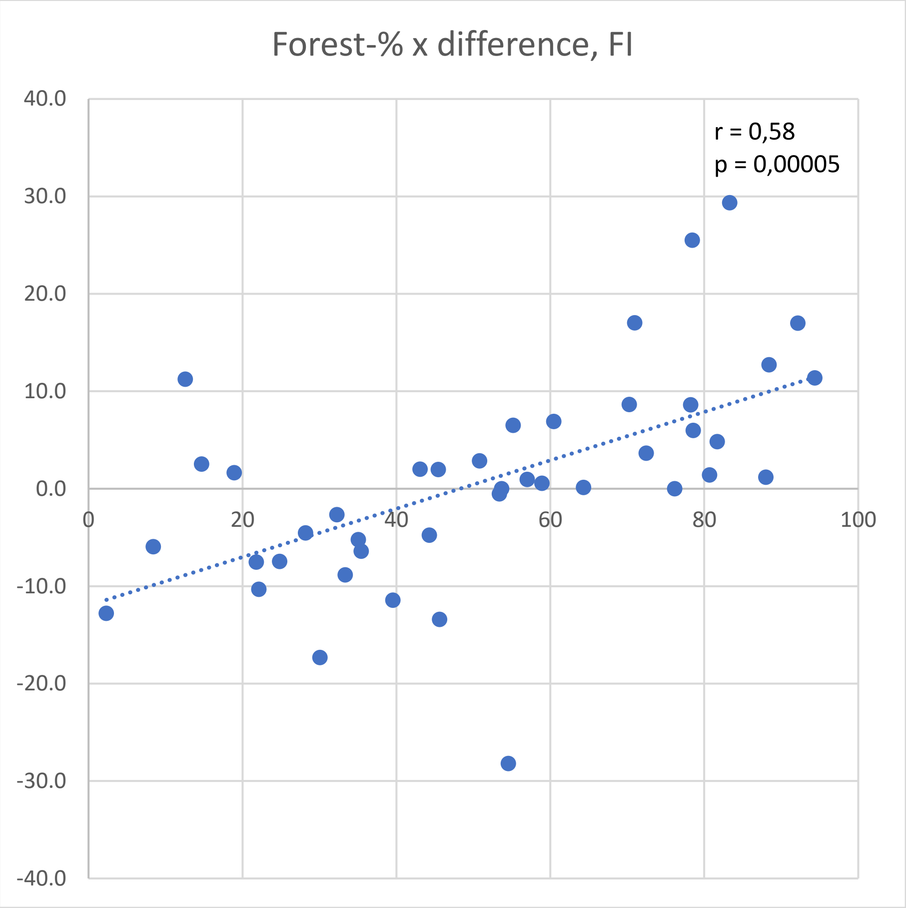



Figure A6.3. Left: Relative effect (%) on individual PA level in the full landscape, measured with the binary fire variables. The black bars represent the 95 % confidence intervals of the estimates. The black star marks the two PAs where the confidence interval for burned area was very large compared to other PAs and variables, so those estimates are displayed separately. Right: Relative effect (%) measured with burned area in the two PAs where the CI was very large.

**Appendix 7**. Pearson correlations between PA forest percentage and the difference between forested areas and the full landscape in the effectiveness estimates derived from the fire variables.



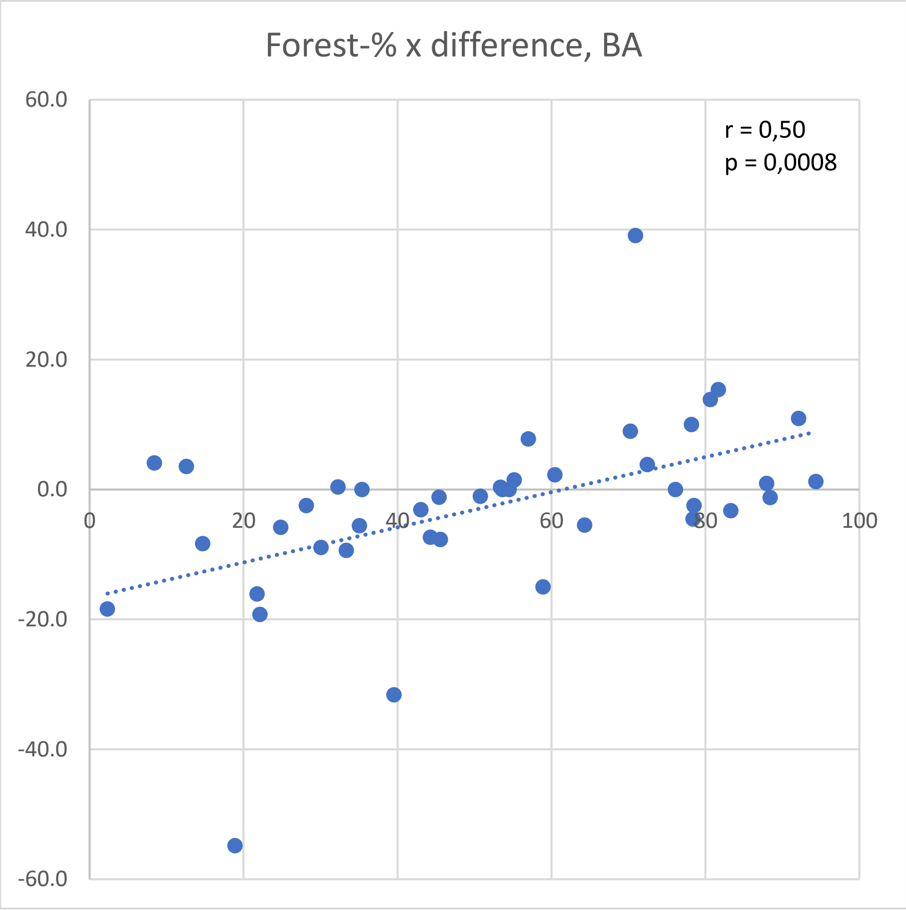




Figure A7. Left: forest-% x difference between forested areas and the full landscape in the effectiveness estimates derived from fire incidence. Right: forest-% x difference between forested areas and the full landscape in the effectiveness estimates derived from burned area. The difference is calculated as “forested areas minus full landscape”.