Analysis of deforestation leakage in tropical forest protected areas using GIS and a global remotely-sensed dataset

By

Robert Whiteside (19048192)

2020

This dissertation is submitted as part of a MSc degree in Environmental Monitoring, Modelling and Management at King's College London.

King’s College London

Department of Geography

MSc Dissertation

I, Robert Whiteside hereby declare (a) that this Dissertation is my own original work and that all source material used is acknowledged therein; (b) that it has been specially prepared for a degree of King’s College London; and (c) that it does not contain any material that has been or will be submitted to the Examiners of this or any other university, or any material that has been or will be submitted for any other examination.

This Dissertation is 8,419 words, sufficient depth could not be achieved within the optional 6,000 word limit.

1st September 2020

Abstract

Protected Areas (PAs) are a key tool in the conservation of biodiversity and ecosystem function. However, PA effectiveness and potential confounding factors must be understood in order to justify the global PA network. This study assesses deforestation rates in newly designated tropical forest PAs using the Global Forest Change dataset; deforestation was quantified within the PA and the surrounding 10km buffer zone before and after designation to identify spillovers that could compromise effectiveness. Statistical matching and difference-in-differences regression was used to generate counterfactual controls and identify significant changes across time, respectively. Designation did not significantly reduce deforestation, although rates were lower than the controls. Leakage was potentially found in 2 of the 9 PAs, however the combination of geo-physical, socioeconomic, and political factors on a local spatial scale require more in-depth analysis for conclusive assessment; further work should be targeted at this scale of inquiry for research on PA spillovers.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BACI – Before-After-Control-Impact
CAM – Cameroon
CBD – Convention on Biological Diversity
COL - Colombia
DiD – Difference in Differences
FAO – Food and Agriculture Organisation
GDP – Gross Domestic Product
GFC – Global Forest Change
GRUMP – Global Rural-Urban Mapping Project
HON - Honduras
HOT – Humanitarian OpenStreetMap Team
IND – India
IQR – Interquartile Range
IUCN – International Union for Conservation of Nature
LOESS – Locally Estimated Scatterplot Smoothing
MEA – Millennium Ecosystem Assessment
MYA - Myanmar
PA – Protected Area
PAE – Protected Area Effectiveness
PER – Peru
PHI – Philippines
PSM – Propensity Score Matching
QQ – Quantile-Quantile
REDD – Reducing Emissions from Deforestation and forest Degradation
SD – Standard Deviation
UN – United Nations
UNEP-WCMC – UN Environment Programme World Conservation Monitoring Centre
WDPA – World Database on Protected Areas

Acknowledgements

I would like to thank my supervisor Professor Nick Drake for all the help and advice, my family and girlfriend for their continued support, and the providers of all the data freely acquired for this project without which it would not have been possible and whose contribution to the collaborative accessibility of science is greatly appreciated.

1. Introduction

The acceleration of technological development and population growth from the 19th century to present has resulted in unprecedented environmental crises (Sanderson *et al*., 2002; Zalasiewicz *et al*., 2010): Climate change, habitat loss, biodiversity declines, and increased extinctions (Dirzo *et al*., 2014; Ceballos *et al*., 2015). As a result, a new geological epoch is proposed – the 'Anthropocene', characterised by humanity's role as the defining global environmental forcing agent, the impacts of which will be recognisable on the scale of geological time (Crutzen, 2002; Zalasiewicz *et al*., 2010). Arresting this degradation of environmental processes is imperative to avoid permanent loss of ecosystem function and the accompanying cost to human society (Sannigrahi *et al*., 2018).

The modern Protected Area (PA) concept was originally conceived in the 19th century to preserve iconic landscapes and wildlife before expanding throughout the 20th century to fill a complex role of managing ecosystem services, supporting local livelihoods, and preserving charismatic flora and fauna (see Figure 1) (Watson *et al*., 2014). PAs are a key tool in the conservation of nature (Dudley, 2008) as evidenced by the expansion of the global PA network (Figure 1), exemplified by the Convention on Biological Diversity's "Aichi target" 11 – to conserve 17% of land and 10% of marine waters by 2020 through PAs and other area-based conservation measures (CBD, 2010); this multilateral treaty, ratified by all UN member states, represents a massive global commitment and currently 15% of land and 8% of oceans fall under a form of protection (IUCN and UNEP-WCMC, 2020). It is critical to the legitimacy of conservation that the restrictions on these vast swathes of land and sea are achieving the desired outcomes.

Although research on Protected Area Effectiveness (PAE) with strong counterfactual study design is relatively sparse (Geldmann *et al*., 2013), it has become clear that simply designating an area 'protected' does not necessarily confer the desired benefits to the environment – there are complex social, economic, and political dynamics that impact effectiveness (Spracklen *et al*., 2015; Pfaff and Robalino, 2017; Fuller *et al*., 2019). Additionally, there is bias in the non-random location of PAs (i.e. remote and inaccessible) and which ecosystems and species are represented (Joppa *et al*., 2008; Joppa and Pfaff, 2011; Watson *et al*., 2014). It is essential that these factors are fully understood and considered

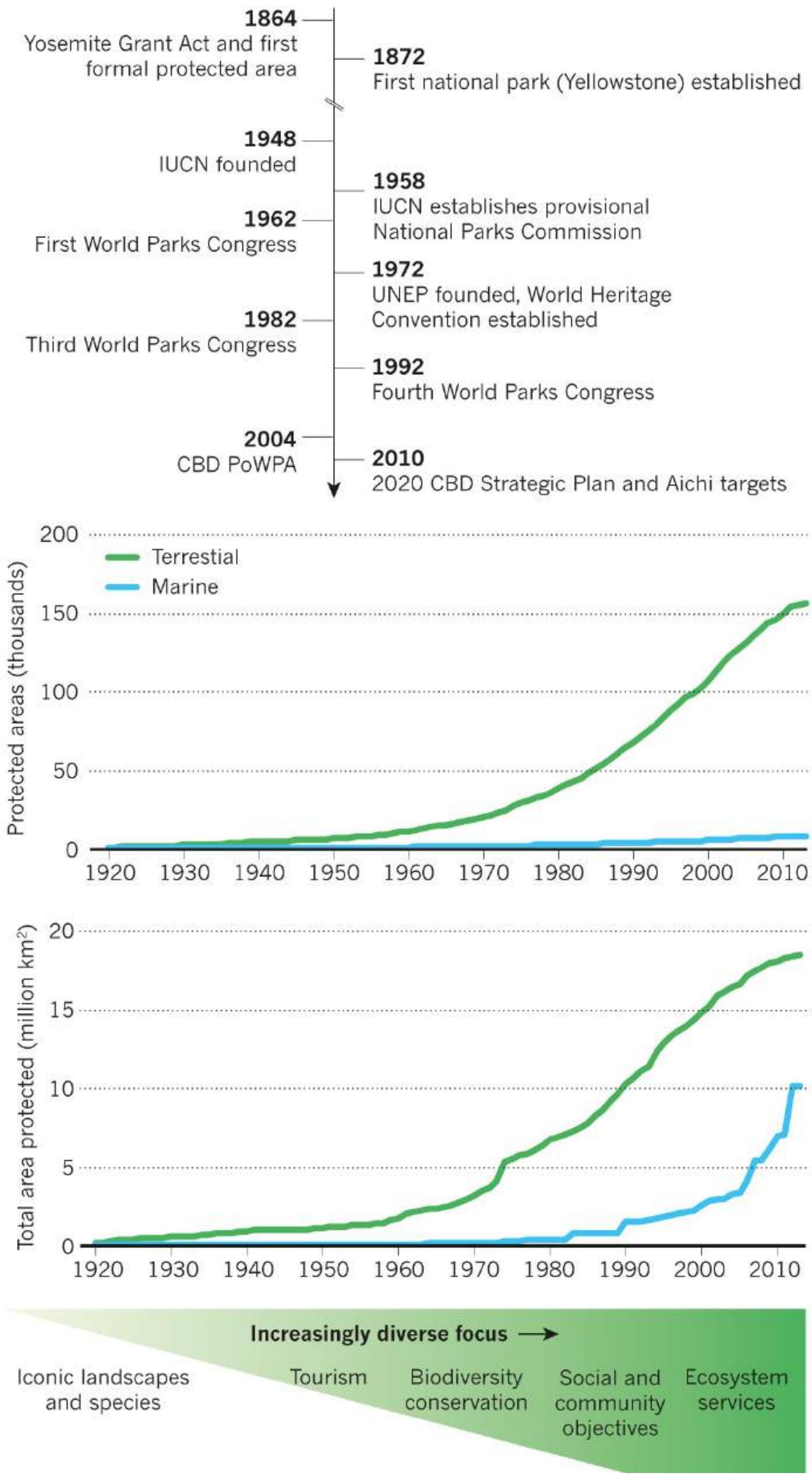

Figure 1. The historical evolution of PAs from the 19th century to present, taken from Watson *et al*. (2014). Top is a timeline of key events and organisations. Middle shows the numerical growth of the number of PAs and the millions of square kilometers covered by the global network. Bottom is the change of the role that protected areas are expected to fill.

when planning and managing PAs, particularly under major resource constraints (Coad *et al.*, 2019). One of the ways that PAE can be inflated is through ‘leakage’: Originally conceptualised regarding carbon sequestration through avoidant deforestation policies (e.g. REDD), whereby net carbon sequestration fails due to demands for timber/land being met elsewhere (Brown *et al.*, 1997; Schwarze *et al.*, 2002; Aukland *et al.*, 2003). In the context of PAs designated for the preservation of ecology/habitats, leakage could result in elevated conversion of land (e.g. deforestation) immediately surrounding the PA; this loss of the ‘buffer zone’ could offset the benefits of the restriction (Ewers and Rodrigues, 2008).

Leakage has been assessed almost exclusively in tropical forest ecosystems; this may be because of the disproportionate representation of the world’s terrestrial biodiversity, severe anthropogenic pressure (Gardner *et al.*, 2010; Giam, 2017), and the recent increase in PA coverage, particularly in South America (Jenkins and Joppa, 2009). Additionally, the loss of ecosystem function through deforestation is one of the most common methods of evaluating PAE (impact evaluation) (Andam *et al.*, 2008; Fuller *et al.*, 2019; Ribas *et al.*, 2020). The loss of buffer zones through leakage is particularly important to identify due to the ecological dependence of tropical PAs on the surrounding ecosystems, especially in terms of maintaining biodiversity and climate change mitigation (DeFries *et al.*, 2005; Laurance *et al.*, 2012; Mitchard, 2018). Although there have been several studies investigating leakage (e.g. Sánchez-Azofeifa *et al.*, 2003; Oliveira *et al.*, 2007; Lui and Coomes, 2016; Poor et al., 2019), there is little consensus on the causes and only Oliveira et al. (2007) investigated the temporal effects after the designation of new land use restrictions.

This study will investigate the temporal effects of PA designation on deforestation using a global remotely-sensed dataset. Counterfactual control samples will be identified using statistical matching to control for bias in PA location. These will be used to compare the deforestation rates of the PA and buffer over time, followed by detailed spatial analysis of the patterns and drivers of deforestation within the buffer zone to identify leakage and potential drivers. The socioeconomic and political status of the regions surrounding the PAs will also be considered to assess commonality between cases and potential future directions for further research.

The research questions for this study are as follows:

Does protected area designation displace deforestation to the surrounding landscape (leakage)?
Are there spatial, socioeconomic, political patterns affecting leakage?

2. Literature Review

2.1 Protected Area Impact Evaluation

The commitments to ecological preservation and the conservation of biodiversity under constrained resources, as outlined in the introduction, necessitate comprehensive and robust evaluation. Impact evaluation in conservation first started gaining attention in the 1990s but was focused on straightforward measurable outputs (e.g. staff trained, km^2 protected, communities instructed etc.) rather than the outcomes (e.g. biodiversity/ecosystem preserved) (Ferraro and Pattanayak, 2006); by the time of the Millennium Ecosystem Assessment in 2005 it was clear that "*Few well-designed empirical analyses assess even the most common biodiversity conservation measures*." (MEA, 2005). This is particularly relevant to PAs as they are broadly judged in terms of their number and area of coverage, with an assumption that effectiveness is inherent (Chape *et al*., 2005), as evidenced by the Aichi targets (see Section 1). The expectation of conservation practitioners, donors, and governing bodies that PAs are justified and demonstrably (with rigorous evaluation) valuable as conservation investments has grown considerably over the past 3 decades (Ferraro and Pattanayak, 2006; Ferraro, 2009; Ferraro and Pressey, 2015; Baylis *et al*., 2016).

In the environmental sciences, establishing an experimental control group is often practically impossible due to ethical, logistic, and financial constraints (Schleicher *et al*., 2020). This creates a challenge of establishing a counterfactual control (i.e. what would have happened with no intervention) in observational studies. Although challenging, in comparison to other fields like economics and public health, the quality of impact evaluation in conservation is poor (Baylis *et al*., 2016). Some studies naively avoid counterfactual thinking and (i) simply compare outcomes of treated against untreated (in this case protected land against unprotected land) (e.g. Rodríguez *et al*., 2013), or (ii) compare outcomes before and after the treatment is implemented (e.g. Macdonald *et al*., 2011) (Ribas *et al*., 2020). However, these methods have assumptions that are unlikely to be upheld in reality – (i) that treatments are randomly selected and distributed, and (ii) that the outcome in question is uniform across time (Joppa and Pfaff, 2010; Baylis *et al*., 2016). For example Lui and Coomes (2016) used control samples within a distance of 25km from the PA boundary, not controlling for confounding variables and assuming that there is a similarity in land characteristics affecting the outcome (deforestation) between the control samples and the PA due to spatial proximity, which has previously been found not to be the case (Joppa and Pfaff, 2011).

Counterfactual thinking requires well-constructed theories of change to determine which characteristics or confounding covariates can affect the conservation outcome in order to control for them (Qiu *et al*., 2018; Schleicher *et al*., 2020). In the case of tropical forest PAs, deforestation is the simplest measure of conservation outcome and used almost exclusively for impact evaluation (Fuller *et al*., 2019), indeed all of the studies referenced herein will be using tropical forest deforestation as a measure of PA evaluation; remote sensing techniques can be used to cover deforestation at a high spatial resolution on a global scale, with data sets now going back decades (Hansen *et al*., 2013). Therefore, the question for the theory of change is: What are the socio-geophysical characteristics (confounding variables) that affect the likelihood of both deforestation and protected area designation, and then how do you control for them? For the former, it is known that both deforestation and PA designation are linked to agricultural suitability and accessibility/remoteness (Andam *et al*., 2008; Venter *et al*., 2018) – PAs are more likely to be established in remote, inaccessible regions where deforestation rates are unlikely to be high and opportunity costs are low; these areas are therefore protected *de facto*, without the necessity of a PA (Andam *et al*., 2008; Joppa *et al*., 2008; Joppa and Pfaff, 2010, 2011; Amin *et al*., 2019). There are a number of techniques used to control for confounding factors, but the most commonly used and effective method in PA impact evaluation is statistical matching (Ribas *et al*., 2020).

2.2 Matching

Matching refers to a range of statistical techniques employed to establish or improve causal inference. It is used across a variety of fields where experimental controls are not feasible, such as economics, medicine, political science, and law (Sekhon, 2011). Controls are selected *ex post* based on a degree of similarity or distance to the treatment group across a range of predefined covariates (Schleicher *et al*., 2020); in this way the aim of matching is to create two sample groups (treatment and control) with similar covariate distributions, resulting in an 'apples to apples' comparison (Joppa *et al*., 2008; Joppa and Pfaff, 2011). The most common methods (Sekhon, 2011) (see Table 1) are nearest neighbor with propensity score matching (PSM) (based on logistic regression) (Rosenbaum and Rubin, 1983), nearest neighbor with Mahalanobis distance (Rubin, 1980), and genetic matching (Diamond and Sekhon, 2013). Method selection should not be carried out *a priori* and the process should be iterative with testing of different methods and models to find the best fit or 'balance' (Schleicher *et al*., 2020), although many of the examples in Table 1 do not show evidence of this process. In a review of PA impact evaluation, Ribas *et al*. (2020) found that studies neglecting

Table 1. Selection of papers' methods and covariates using matching for PA impact evaluation.

Paper	Matching Method	Matching covariates
Andam et al. (2008)	Nearest-neighbor using the Mahalanobis distance metric, with replacement, 0.5SD caliper	Distance to road, distance to city, distance to forest edge, land class (soil fertility, slope, humidity)
Gaveau et al. (2009)	Propensity score matching, without replacement	Elevation, slope distance to forest edge, distance to road
Joppa and Pfaff (2011)	Nearest-neighbor using the Mahalanobis distance metric, without replacement	Elevation, slope, ecoregion, distance to road, distance to city, agricultural suitability
Nelson et al. (2011)	Nearest-neighbor using the Mahalanobis distance metric, with replacement, 0.5SD caliper	Elevation, slope, ecoregion, distance to roads, distance to city, administrative region
Alix-Garcia et al. (2012)	Nearest-neighbor using the Mahalanobis distance metric, with replacement	PA area, slope, elevation, baseline forest type, prior deforestation rates, municipal population density, degree of marginality, and access to markets.
Rasolofoson et al. (2015)	Nearest-neighbor using the Mahalanobis distance metric	Elevation, slope, distance to forest edge, distance to road, distance to recent deforestation, agricultural suitability, population density
Anderson et al. (2018)	Propensity score matching, with replacement	Elevation, slope, distance to road, distance to city, distance to river, agricultural suitability, geopolitical state
Geldman et al. (2019)	Propensity score matching, without replacement, 0.25SD caliper	Elevation, slope, access, temperature, precipitation, human footprint, land cover, soil type, nutrient levels
Herrera et al. (2019)	Combined propensity score matching with nearest-neighbor using the Mahalanobis distance metric	Distance to road, slope, distance to forest edge, soil quality, rainfall, vegetation
Oldekop et al. (2019)	Propensity score matching (full matching)	PA area, baseline forest cover, poverty, elevation, slope, precipitation, population density, agricultural effort, international migration, time to travel to population centres
Poor et al. (2019)	Generalized boosted regression models (modified genetic matching)	Elevation, slope, distance to road, distance to city, distance to forest edge, presence or absence of forest in 2016
Yang et al. (2019)	Propensity score matching, with replacement, 0.5SD caliper	Elevation, slope, tree cover, distance to forest edge, aspect, terrain roughness, topographic wetness index, human influence index, travel time to the nearest city, precipitation, temperature, soil carbon, soil depth, soil acidity, amount of bulk and clay in the soil

counterfactual study design typically over-estimate PAE; confirming the findings of the first implementations of matching in PA impact evaluation by Andam *et al*. (2008) and Joppa and Pfaff (2011). In general, matching studies have found that PAs do confer benefits to tropical forest ecosystems but these can be marginal and must not be assumed (Andam *et al*., 2008; Gaveau *et al*., 2009; Joppa and Pfaff, 2011; Rasolofoson *et al*., 2015; Geldmann *et al*., 2019).

2.3 Leakage

Leakage is a form of spillover, whereby effects from a conservation intervention are displaced onto non-intervention areas; this can be negative (leakage), or positive, known as 'blockage' – such as ecological benefits of population reservoirs within proximal PAs (Ewers and Rodrigues, 2008; Fuller *et al*., 2019). Spillovers can occur across a range of spatial scales, for example conservation interventions might raise the cost of timber in one region causing increased deforestation in another distant region with lower costs (Moilanen and Laitila, 2016). This 'indirect' leakage can involve complex market dynamics and occur across national boundaries, such as Thailand's 1989 logging ban increasing deforestation in Cambodia and Myanmar (Gan and McCarl, 2007; Henders and Ostwald, 2014; Lim *et al*., 2017; Pfaff and Robalino, 2017). For the purpose of this study however, 'neighbourhood' (or 'direct') leakage will be exclusively examined, whereby land-use conversion and deforestation is offset to the immediate surroundings or 'buffer zone' of a PA.

Not accounting for spillovers in PA impact evaluation could result in significant inaccuracy, especially due to the aforementioned bias in PA designation for land with low opportunity costs rather than high biodiversity – human pressure could in theory be displaced onto areas with a higher ecological value (Venter *et al*., 2018). Buffers are also important to preserve because they maintain ecological health by increasing species capacity and connectivity of habitats (Sayer, 1991; Bennett and Mulongoy, 2006; DeFries *et al*., 2005) and loss of buffer will result in degradation within the PA as found by Curran *et al*. (2004) in Kalimantan. Buffer zones are also experiencing elevated population growth in Africa and Latin America (Wittemyer *et al*., 2008) and even in remote regions deforestation pressure is still present (Fuller *et al*., 2019).

There have been a number of PA impact evaluations that have included a buffer zone in their analysis, although not always for the purpose of assessing spillovers (e.g. Sánchez-Azofeifa *et al*., 2003; Curran *et al*., 2004; DeFries *et al*., 2005). Global-level studies have found leakage only in a small number of PAs and blockage as a much more common spillover (Joppa and Pfaff, 2011; Lui and Coomes, 2016; Fuller *et al*., 2019); although only Joppa and Pfaff (2011) used robust counterfactual matching and spillovers was not the focus of their study. Fuller *et al*.'s (2019) global meta-analysis found that national-scale socioeconomic factors (population growth, proportion of agricultural land, and forestry product value) had potential as drivers of spillover. However, because patterns of spillover have generally not been found across different nationalities or regions this seems unlikely (Joppa and Pfaff, 2011; Lui and Coomes, 2016). For example, Poor *et al*.'s (2019) counterfactual assessment of Sumatran PAs found varying degrees of positive and negative spillover, as did Robalino *et al*. (2017) in Costa Rica and Herrera *et al*. (2019) in Brazil; if spillovers can vary significantly within the same country, is it valuable to try and average spillover effects regionally or globally in an attempt to draw out universal, large-scale drivers as Joppa and Pfaff (2011) and Fuller *et al*., (2019) among others, have done? This inevitably dilutes the effects of what is clearly a relatively uncommon issue, suggesting insignificance. Robalino *et al*. (2017) found that leakage is directly related to distance to roads and from the PA entrances, supporting the findings from theoretical and modelling studies that heterogeneous local factors are the key drivers of spillovers such as policy, management, infrastructure, workforce mobility, and tourism (Bode *et al*., 2015; Renwick *et al*., 2015; Delacote *et al*., 2016; Pfaff and Robalino, 2017; Amin *et al*., 2019).

There is a major gap in the PA spillover literature regarding the effect immediately following PA designation. Oliveira *et al*. (2007) found high neighbourhood leakage in the Peruvian Amazon following logging concessions, with deforestation rates increasing up to 400% (without counterfactual controls). Even though Ewers and Rodrigues (2008) drew attention to this study in their oft-cited paper, to my knowledge there has not been a before-after-control-impact (BACI) study of deforestation spillovers from PAs. This may be because of the difficulties in establishing a counterfactual baseline deforestation rate and avoiding the assumption that rates are stable over time as criticised by Joppa and Pfaff (2010) and Ribas *et al*. (2020). Additionally, both Joppa and Pfaff (2011) and Fuller *et al*. (2019) suggest that there is a lag in land-use change so PAE assessments should focus on PAs established long before the deforestation data begins; however, these statements are not justified or drawn from their empirical findings. A regression-based technique that has been used in deforestation BACI experiments is 'difference-in-differences' (DiD) – used for longitudinal data to estimate causal effect when both treatment and control outcomes are known over time (Lechner, 2011); DiD was used by Prem *et al*. (2020) for post-conflict Colombian deforestation and in combination with matching for PA impact evaluation by Shah and Baylis (2015) and Anderson *et al*. (2018).

This study will attempt to address the research aims in Section 1 and the gaps in the literature by performing a counterfactual assessment of spillover from newly designated PAs using a combination of matching and DiD; possible cases of leakage will be investigated further with fine-scale spatial analysis.

3. Materials & Methods

3.1 Data

The Global Forest Change (GFC) dataset was published by Hansen *et al*. (2013) and subsequently updated every year; the data are freely available for download and use under a Creative Commons Attribution 4.0 International License and have been cited over 3500 times according to Web of Science. A classification method adapted from Potapov *et al*. (2012) was applied using Google Earth Engine to 654,178 Landsat 7 ETM+ growing season images to classify forest (≥50% canopy cover). Forest loss was defined as "*a stand-replacement disturbance or the complete removal of tree cover canopy at the Landsat pixel scale* [30x30m]". The use of high resolution (30m pixels) Landsat 7 means that small scale disturbances are still identified, which can be significant in habitat loss, especially when adding to existing clearings (Ryan *et al*., 2012). Furthermore, fine-scale differences in deforestation, for example at PA boundaries, will need to be captured accurately for this analysis. The dataset consists of a number of forest metrics covering global land surfaces (see Table 2 and Figure 2), of which the tree canopy cover and forest loss layers will be used in this study.

Table 2. Data layers available from Hansen *et al*.'s (2013) Global Forest Change.

Name	Description
Tree canopy cover (year 2000)	Percentage of pixel covered by canopy of >5m vegetation in 2000
Forest cover gain	Conversion of non-forest to forest for the period 2000-2019 (1 = gain or 0 = no gain)
Year of forest loss	Year of conversion from forest to non-forest (0 for no conversion, 1-19 for conversion in years from 2001-2019 respectively)
Data mask	No data (0), land (1), permanent water body (2)
First available reference Landsat 7 multispectral image	The first available (typically year 2000) cloud free Landsat composite image
Last available Landsat 7 multispectral image	The last available (typically year 2019) cloud free Landsat composite image

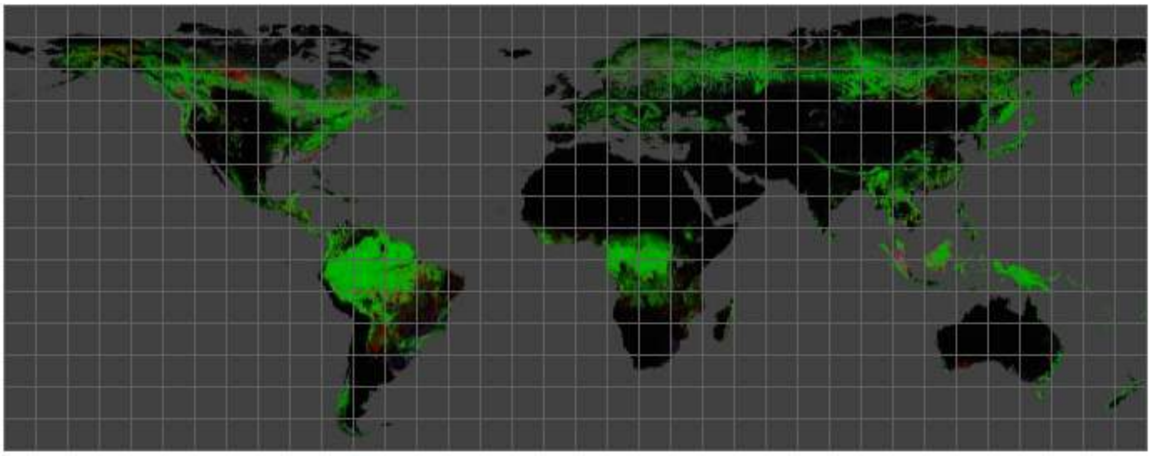

Figure 2. Map of the global forest change dataset (Hansen *et al*., 2013). Each tile is a 10x10 degrees downloadable unit.

The World Database on Protected Areas (WPDA) (UNEP-WCMC and IUCN, 2020a) is the most comprehensive catalogue of the global PA network, constantly being updated and (as of August 2020) containing freely accessible data on 261,766 PAs (Bingham *et al*., 2019; UNEP-WCMC, 2020). These data include the spatial boundaries of the PAs accompanied by 29 meta-data attributes, such as date and type of designation etc. (see Figure 3) (UNEP-WCMC, 2019). The WDPA was subsetted based on 6 criteria shown in Table 3 to identify appropriate PAs.

The WWF Terrestrial ecoregions of the world dataset is a freely available GIS compatible map of 867 ecoregions that cover the global landmass, developed in collaboration with *"over 1000 biogeographers, taxonomists, conservation biologists, and ecologists from around the world"* (Olson *et al*., 2001). This map was designed with biodiversity and conservation planning at its core and delineates clear boundaries useful for fine-scale spatial analyses (Liu *et al*., 2018). These data were used to further subset the WDPA by manually checking of each PA to ensure that it was protecting a tropical or sub-tropical forest ecoregion and also used in the sampling structure (see Section 3.2). The final selection of suitable PAs (n=9) can be seen in Table 4.

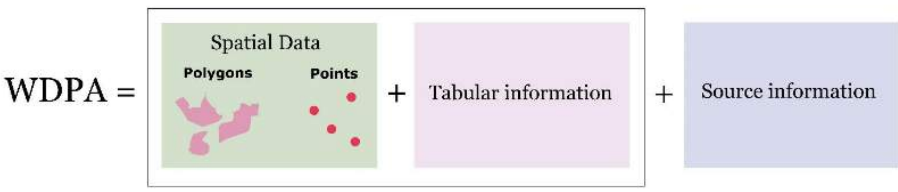

Figure 3. Structure of the WDPA, taken from user manual (version 1.6) (UNEP-WCMC, 2019).

Table 3. Criteria and justification for subsetting the WPDA

Subset	Justification
Located in the tropics (between 23.43661°N and 23.43661°S).	This study is looking at the impact of PAs on tropical forest ecosystems.
Terrestrial.	Marine PAs are not relevant.
Designated status.	The WDPA contains proposed PAs, which would not have a measurable effect.
2006 ≤ Designation date ≤ 2014.	To be within the deforestation data (GFC) range (2001–2019) and have at least 5 years before and after designation (see Appendix A for comparative study lengths).
IUCN category I, II, or IV.	These are strict no-take PAs, ensuring that sustainable-use is not misattributed to PA failure.
>20km from any other PAs.	To avoid overlap of PA 10km buffer zones that would confound estimates of individual PA spillover.

The remaining data relates to the confounding covariates used in the matching analyses; following the theory of change for PA deforestation covered in Section 2, the covariates were selected as quantifiable measures of remoteness and low opportunity cost: Tree canopy coverage (%), distance from roads (m), distance from human settlements (m), distance from forest edge (m), elevation (m), and slope (%). This selection is supported by the representation of these covariates in the published studies in Table 1. A full breakdown of the data layers, dates, and accessibility is available in Appendix B. Administrative areas and roads were accessed either from the respective country's government cartographic/statistical department if available, or from the Humanitarian OpenStreetMap Team (HOT), an open source non-profit mapping organisation. OpenStreetMap is the largest and most successful crowdsourced geospatial data project (Minghini and Frassinelli, 2019) and despite accuracy concerns with open source data, in general this has not been found to be an issue (Zhang and Malczewski, 2017; Nasiri *et al.*, 2018). Human settlements were acquired from the Global Rural-Urban Mapping Project (GRUMP), a global dataset derived from year 2000 night-lights following the method of Balk *et al.* (2006). Elevation was provided by NASA's Shuttle Radar Topography Mission (SRTM) digital elevation model (USGS, 2014), offering global coverage at 30m resolution, with an accuracy of ±7-9m (Rodriguez *et al*., 2006), adequate for the purpose of this study.

Table 4. Final Selection of PAs from WDPA, their location, designation date, and size.

Name	**WDPA ID**	**Country**	**Year Designated**	**Size (km^2)**
Boumba Bek/Nki	308624 & 30674	Cameroon (CAM)	2005	2361.76 / 3129.65
Deng Deng	555547995	Cameroon (CAM)	2013	687.35
San Miguel de los Farallones	555555800	Colombia (COL)	2011	33.79
Congolón, Piedra Parada y Coyocutena	62051	Honduras (HON)	2010	110.46
Montaña de Botaderos Carlos Escaleras Mejía	555582981	Honduras (HON)	2012	967.55
Papikonda	1774	India (IND)	2008	1012.86
Kyauk Pan Taung	1235	Myanmar (MYA)	2013	130.6
Bosques Nublados de Udima	555544103	Peru (PER)	2011	1218.32
Mount Balatukan Range	555583087	Phillipines (PHI)	2007	84.23

3.2 GIS Processing

All processing and management of the spatial data was carried out in QGIS v3.4 (QGIS Development Team, 2020), see Figure 4 for the full workflow.

In order to have distances in consistent standard units (metres), all spatial data was reprojected to a suitable projected coordinate system (see Appendix C for details) (Longley *et al*., 2015). Roads and GRUMP human settlement layers were both converted to raster format and a proximity grid was generated, giving a distance to the nearest road and settlement in metres for every pixel (30m resolution to match the GFC data). A binary forest classification layer was generated by classifying all pixels ≥50% tree canopy cover as forest using the raster calculator on the GFC tree canopy cover data (note that 50% is following the classification of the GFC (Hansen *et al*., 2013) but greatly exceeds the FAO's 10% canopy cover definition (FAO, 2020a)). Creating a proximity grid on the binary forest layer (targeting non-forest) generated a distance to the forest edge layer. A binary deforestation layer was also created using the GFC loss year data. Slope was calculated using QGIS's slope command on the SRTM elevation data.

A buffer of 10km from the PA boundaries (taken from the WDPA) was created; this is a relatively arbitrary distance and results are dependent on the distance used (DeFries *et al*., 2005), however 10km appears to be the standard for other studies (Lui and Coomes, 2016; Poor *et al*., 2019) (see Table 5).

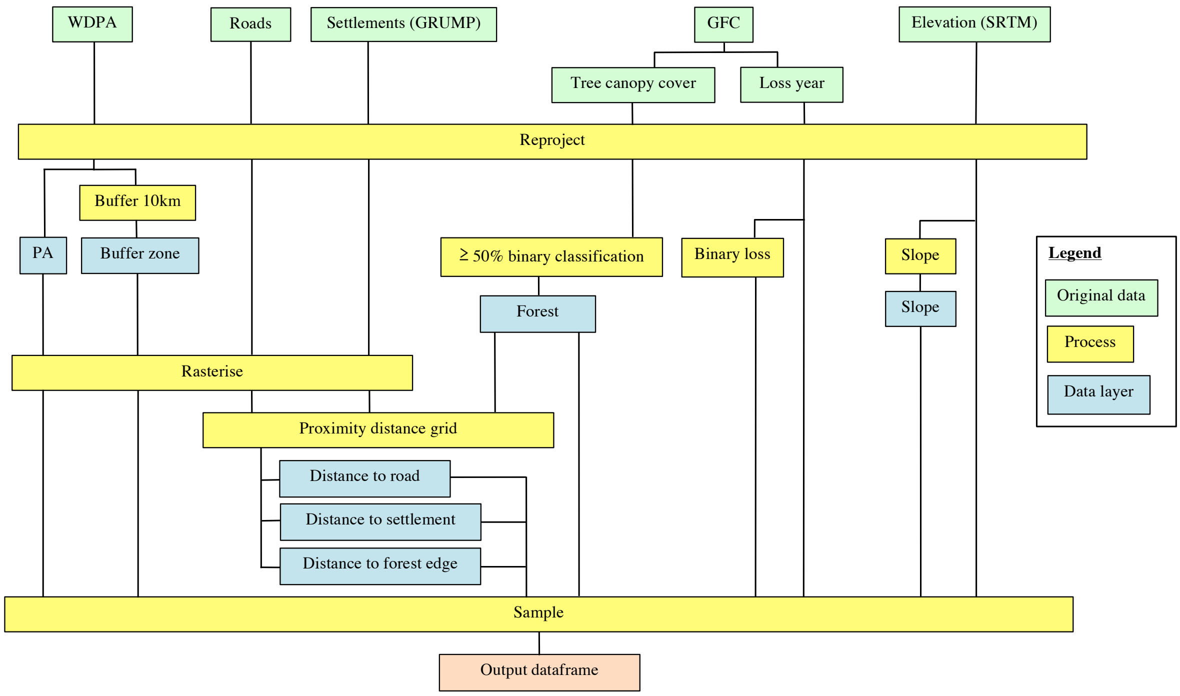

Figure 4. Flowchart of the GIS workflow used to produce the sample for buffer and PA matching.

Control sample points for the PA and buffer were randomly generated with the following pixel requirements: Matching administrative region (avoiding confounding administrative policy differences), matching ecoregion, outside of other PAs and buffer zones, and forested (≥50% tree canopy cover in year 2000). Treatment sample points were randomly generated within forested pixels of the PA and then buffer. To prevent replication, a minimum distance of 30m (1 pixel) was required. The buffer and PA were sampled separately with unique control samples for each as recommended by Negret *et al*. (2020). Sample size was determined through trials of analyses to maximise the size and statistical power within computational limits as done by Rasolofoson *et al*. (2015), with the aim of having 2-4 times control samples to treatment samples as in Rasolofoson *et al*. (2015), Joppa and Pfaff (2011), and Andam *et al*. (2008); this resulted in 5,000-10,000 treatment samples and 30,000-60,000 samples, depending on the size of the region covered. Using the QGIS plugin ‘Point sampling tool’ (Jurgiel, 2020), data was extracted from the following layers at each sampling point: Deforestation year, binary deforestation, tree canopy cover, elevation, slope, distance to road, distance to settlement, distance to forest edge, and an additionally binary layer denoting treatment (1) or control (0).

Table 5. Selection of PA impact evaluation studies and the distance from the PA boundary considered within the buffer zone.

Paper	**Buffer zone (distance from PA)**
Sanchez-Azofeifa et al. (2003)	1 & 10km
Curran et al. (2003)	10km
Oliveira et al. (2007)	20km
Andam et al. (2008)	0-8km (2km intervals)
Armenteras et al. (2009)	10km
Gaveau et al. (2009)	10km
Joppa and Pfaff (2011)	10km
Rodriguez et al. (2013)	2.5 & 5km
Spracklen et al. (2015)	15km (1km intervals)
Lui and Coombes (2016)	0-10km (1km intervals)
Fuller et al. (2019)	1, 2, 5, & 10km
Poor et al. (2019)	10km

3.3 Matching

Matching, post-matching analysis, data manipulation, and visualisation were performed using R v3.6.3 (R Core Team, 2020) in the R studio environment v1.1.463 (R Studio Team, 2016). As emphasised by Sekhon (2011) and Schleicher *et al*. (2020), matching methodology should not be determined *a priori* but through iterative testing with the data. High quality matching results in achieving ‘balance’ between the covariate distributions of the control against the treatment samples, determined by the difference in standardised means (ideally <0.1, but <0.25 is acceptable (Stuart, 2010)) and visual assessment of quantile-quantile (QQ) plots and histograms (Stuart, 2010; Sekhon, 2011; Schleicher *et al*., 2020); QQ plots should show the matched covariate distributions of treatment against control lying on a straight line of $y = x$ through the origin, histogram distributions should match in shape.

Using the MatchIt package v3.0.2 (Ho *et al*., 2007; 2011) three common methods of matching were tested on each PA: Genetic matching (Diamond and Sekhon, 2013) had very high processing times and produced a small sample of matched data (n<1000); nearest neighbour matching with the Mahalanobis distance (Rubin, 1980) achieved inferior balance in comparison to nearest neighbour PSM (Rosenbaum and Rubin, 1983), as the example in Table 6 and Figure 5 shows. For these reasons PSM was selected; further trials determined that the optimal configuration for the PSM was without replacement (each treatment matched to a single control) and with a caliper of 0.1SD (any control >0.1 standard deviations from the treatment was excluded).

Table 6. Results of Mahalanobis and PSM on Bosques Nublados de Udima (PER). Note the standardised mean difference between the two methods, highlighted in bold.

Covariate	**Before matching**			**After PSM matching**			**After Mahalanobis matching**		
	Means Treated	Means Control	Std. Mean Diff.	Means Treated	Means Control	Std. Mean Diff.	Means Treated	Means Control	Std. Mean Diff.
Distance to settlement	2449.65	1836.28	**0.57**	2341.38	2355.24	**-0.01**	2449.65	1930.37	**0.49**
Distance to road	1024.85	1055.80	**-0.05**	983.10	1004.25	**-0.03**	1024.85	981.35	**0.06**
Slope	49.13	50.44	**-0.05**	47.54	47.24	**0.01**	49.13	49.65	**-0.02**
Elevation	2304.25	2760.81	**-0.64**	2681.65	2706.97	**-0.04**	2304.25	2696.40	**-0.55**
Distance to edge	389.39	66.42	**1.04**	136.56	126.07	**0.03**	389.39	85.92	**0.98**
Tree canopy cover	81.54	54.15	**1.38**	71.38	70.03	**0.07**	81.54	64.56	**0.85**

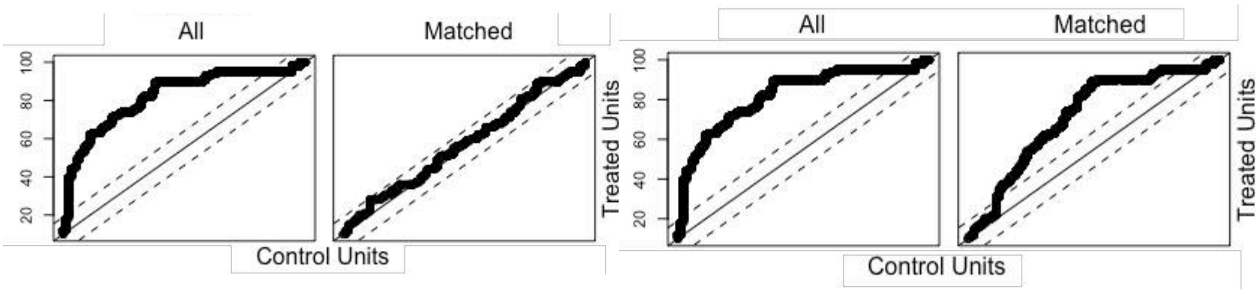

Figure 5. QQ plots of the tree canopy covariate of the Peruvian PA Bosques Nublados de Udima before and after matching between control and treatment samples. Left is from PSM, right is Mahalanobis matching.

PSM uses a logistic regression to generate a propensity score, where the dependent variable acts as an indicator of treatment or control modelled against the covariates (Sekhon, 2011) – the propensity score is a combined probability of the sample receiving treatment. For this study, balance was not achieved using the raw covariates, so as Sekhon (2011) advises, second order polynomials of each covariate were added to the model, significantly improving balance by reducing non-linearity. The MatchIt package then uses a greedy nearest neighbour algorithm to locate the control with the greatest similarity of propensity score to each treatment sample (Ho *et al*., 2011).

3.4 Post-matching analysis

The pre- and post-designation deforestation rates (% yr^{-1}) for the control and treatment groups of the PA and the buffer were assessed visually and Mann-Whitney U tests were used to determine significant differences between groups (e.g. treatment before *vs* treatment after, buffer before *vs* buffer after, treatment before *vs* buffer before etc.). To determine significance of change over time a DiD regression model was used to establish if the deforestation rates between the treatment and the control diverge after PA designation. The key assumption of DiD is of ‘parallel trends’: Without intervention of the PA, the treatment and control groups would have the same trend of deforestation rate over time. The use of matching in this study to produce an ‘apples to apples’ comparison attempted to fulfil this assumption. Following Angrist and Pischke (2008) and Prem *et al*. (2020), Equation 1 is the model used:

$$y = \beta_0 + \beta_1 D^{ti} + \beta_2 D^{tr} + \beta_3 (D^{ti} \times D^{tr}) + \varepsilon \quad (1)$$

Where:

y = Outcome of interest (deforestation

β_0 = Intercept

$\beta_1 D^{ti}$ = Dummy variable of before and after treatment

$\beta_2 D^{tr}$ = Treatment/control dummy variable

$\beta_3(D^{ti} \times D^{tr})$ = Interaction variable of the two dummy variables

ε = Residual error

If the β_3 interaction variable is found to be significant, this suggests that the change in trend observed after the treatment takes effect is independent of the control trend. For the buffer zone, if this significant trend is elevated deforestation, this suggests that leakage has occurred. Model validity was assessed using F-tests, r^2, and residual plots: Fitted *vs* residuals, normal QQ, scale-location, and residuals *vs* leverage.

Further spatial analysis of deforestation was performed on the PAs with potential leakage, as in Spracklen *et al.* (2015) and Lui and Coomes (2016): Deforestation rates were calculated within 1km concentric rings extending from the PA boundary both inward and outward, allowing fine resolution analysis of the spatial trends within the PA and the buffer zone. The inner rings extend as far into the PA as possible and the outer rings will extend beyond the buffer (10km) to 15km from the PA boundary. These were generated using QGIS and the deforestation rates within each ring extracted directly using the QGIS Semi-Automatic Classification Plugin (Congedo, 2018) on the GFC loss layer. Data was imported into R, plotted using a LOESS (Locally Estimated Scatterplot Smoothing) curve, and visually interpreted. If leakage is occurring, it is expected that post-designation deforestation will increase closer to the boundary within the PA and peak within the buffer, exceeding the rates found in the control group (Spracklen *et al.*, 2015; Lui and Coomes, 2016).

To reveal the drivers of deforestation in the buffer zones and how they change over time, human presence and transport links were modelled against probability of deforestation, using logistic regression for the pre- and post-designation time periods; this is appropriate due to the binary outcome variable of deforestation (presence or absence) (Hosmer *et al.*, 2013a). Due to visual interpretation of the deforestation data overlaid onto satellite imagery, distance to river was included as a potential driver; rivers were digitised in QGIS using high-resolution satellite imagery (Google 2020a; 2020b). Distance to river, settlement, road, and

PA boundary rasters were generated following the same methods as previous and randomly sampled with the GFC loss data using 5000 points within the 10km buffer zone.

The sample was imported into R and logistic regression was performed; the coefficient goodness of fit was assessed using an ANOVA of residual deviance, model predictive accuracy was assessed by calculating the misclassification error, and overall model evaluation was performed using the likelihood ratio test (Peng *et al.*, 2002; Hosmer *et al.*, 2013b; Fox and Weisberg, 2019). The model can be represented as follows in Equation 2 (Peng et al., 2002):

$$logit(Y) = \ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \qquad (2)$$

Where:

Y = Outcome of interest (deforestation)

π = Probability of outcome event

α = Intercept

β_x = Regression coefficients

X_x = Predictors (Distance to river, settlement, road, and PA boundary)

4. Results

4.1 Matching

The mean number of successfully matched samples was 2851 for the PAs and 5098 for the buffer zones. Overall, balance was achieved moderately well, with a few exceptions (a full breakdown of the difference in standardised mean and covariate QQ plots and histograms is available in Appendix D and E, respectively). Standardised mean difference of the covariates between treatment and control samples post-matching was excellent, with 87% ≤0.1 and 97% ≤0.25. The QQ plots of the covariate distributions before and after matching all show improvement, but were not perfect post-matching – in particular distance to settlement and roads did not balance optimally, often departing from the desired straight line through the origin.

There were a number of additional issues: San Miguel de los Farallones (COL) PA achieved very poor balance (across the standardised means and QQ plots) with a small matched sample size (<1000); as a result the PA control sample was not included in the following analyses. Congolón, Piedra Parada y Coyocutena (HON) PA also did not achieve good balance but removing elevation improved the model significantly. Finally, removal of distance to settlement from the model of the Mount Balatukan Range (PHI) PA and buffer was required to achieve satisfactory balance, probably due to the low number of settlements on the island.

4.2 Deforestation rates

The observed deforestation rates (% yr^{-1}) range from 0.00-4.40, with the majority of the means falling between 0.1-0.8 (see Appendix F for the means of each treatment). As can be seen in Figures 6 and 7, deforestation rates generally increased between the pre-designation and post-designation time periods, with few reductions and but varying degrees of significance (Mann-Whitney U $p<0.05$) between different treatment groups (see Appendix G for full significance testing between treatment groups). The most consistent significant change was in the PA group with 55.6% experiencing significant increase, compared to 44.4% of the buffer group. The control groups were more varied in their change over time, with less consensus in direction; however 55.6% of the PA control group's pre-designation rates are significantly elevated in comparison to the PA, this drops to 44.4% of cases in the post-designation period.

The buffer zones generally had higher deforestation than the PAs, however only 33.3% were significantly so (for both time periods). The buffer control groups were also elevated in comparison with the buffers, 44.4% significantly different in both time periods. Notable individual cases include the considerable increase within the PA of Montaña de Botaderos Carlos Escaleras Mejía (HON) and the buffer of Kyauk Pan Taung (MYA) post-designation.

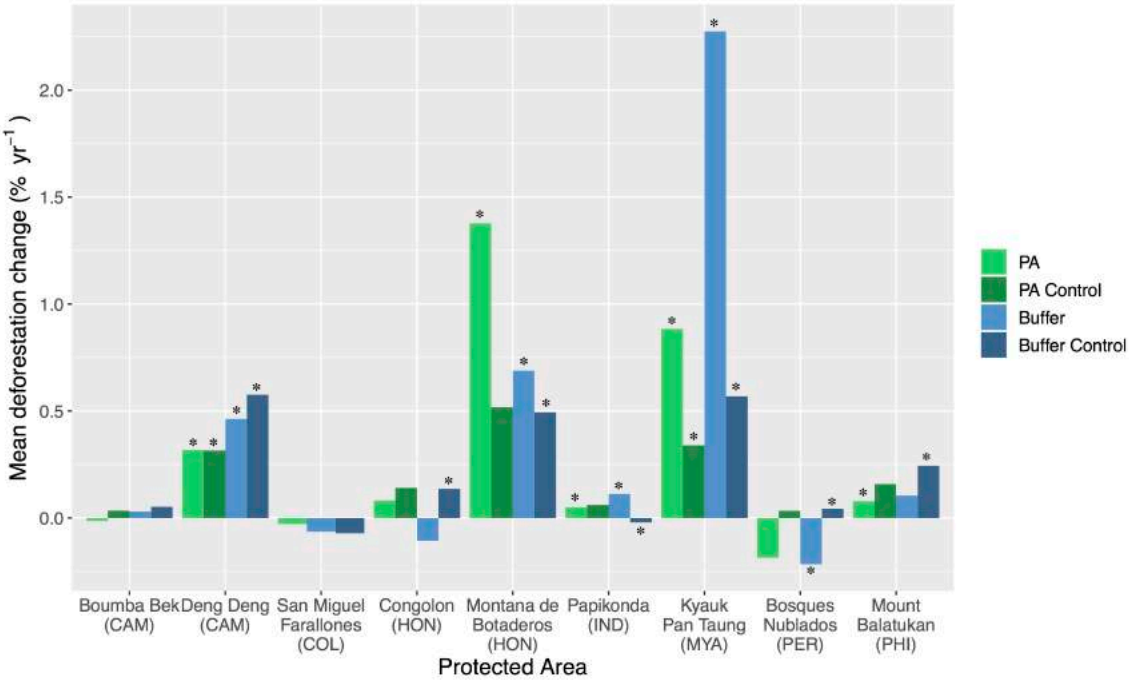

Figure 6. The change in mean deforestation rate between pre- and post-designation of PA in the PA, PA control, buffer zone, and buffer zone control. Significant (Mann-Whitney U test p<0.05) change denoted by *.

4.3 Difference-in-differences

Assessing the raw rates and change in mean deforestation can mask trends over time, making comparisons between treatment and controls difficult; the DiD analysis attempts to account for this. As can be seen in Table 7, the DiD linear models all had significant F-tests apart from the buffer of Boumba Bek/Nki (CAM) and both models of Bosques Nublados de Udima (PER), however the r^2 values were very varied, most models failing to account for >50% of the variation. Assessing the residual plots (see Appendix H) casts doubt on the validity of most of the models, only the buffers of San Miguel de los Farallones (COL) and Papikonda (IND) could be judged to not breach the linearity and homoscedasticity assumptions.

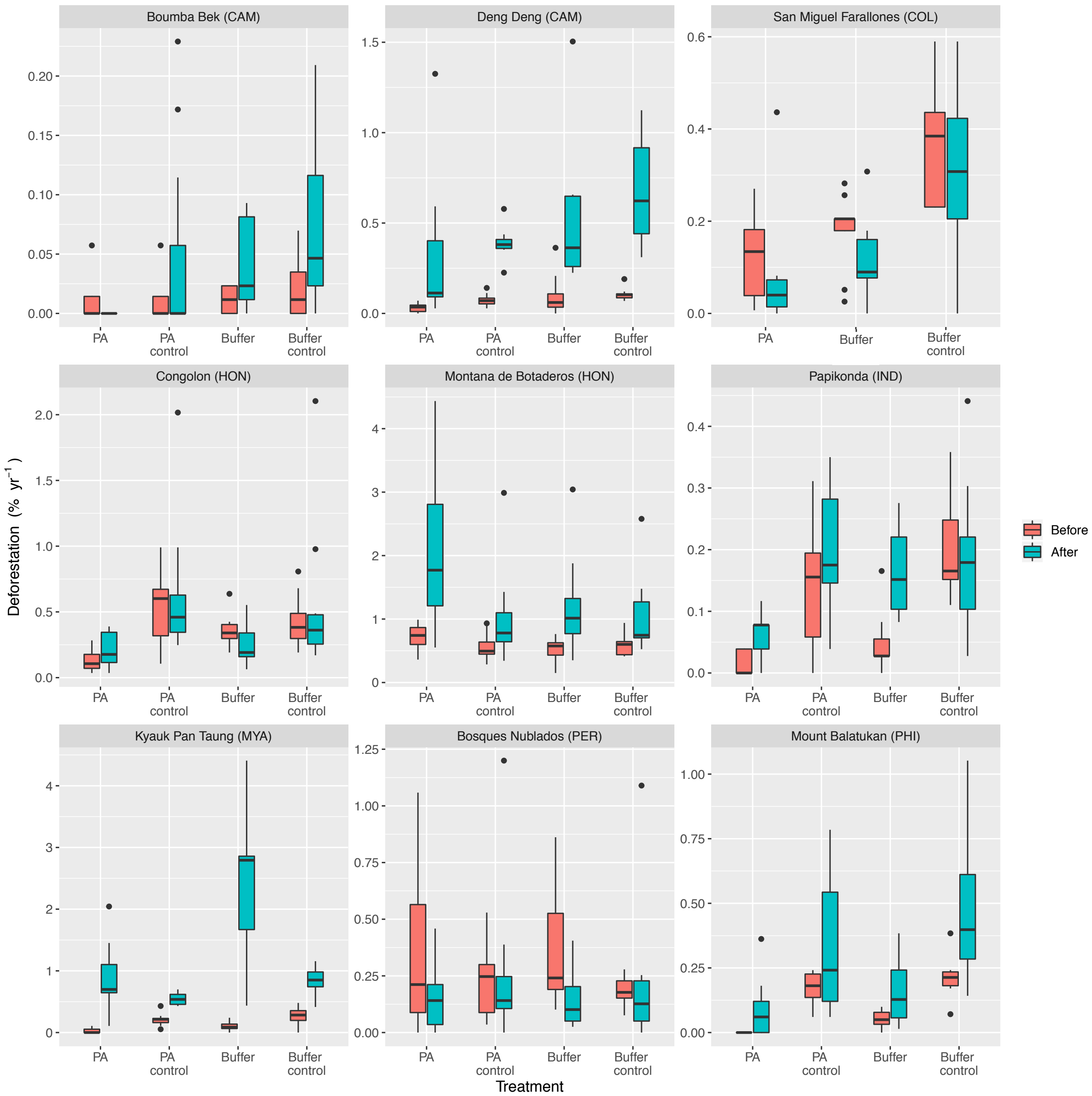

Figure 7. Boxplots of the deforestation rates of the 9 PAs for all treatment types before and after designation (2001-2019). Boxes represent the interquartile range (IQR) with the median value a horizontal line, the minimum and maximum range is shown by the 'whiskers' and outliers (defined as >1.5IQR) marked by black dots.

The significance of the coefficient of interest $\beta_3(D^{ti} \times D^{tr})$ for the buffers of Papikonda (IND) and Kyauk Pan Taung (MYA) suggests that there is a treatment effect independent of the trend of the control groups. As can be seen in Figure 8, this effect is elevated deforestation following PA designation, suggesting that leakage has occurred. In Kyauk Pan Taung (MYA) this is also seen to a lesser extent within the PA, deforestation increasing far above the control group following designation.

Table 7. Results of the linear DiD models for each PA and buffer zone. P-value significance is denoted by * (≤0.05), ** (≤0.01), *** (≤0.001).

Protected Area	**Treatment**	***F* (3,34)**	r^2	β_0	$\beta_1 D^{ti}$	$\beta_2 D^{tr}$	$\beta_3(D^{ti} \times D^{tr})$
Boumba Bek/Nki (CAM)	PA	2.85*	0.20	0.01	0.00	0.04	-0.05
	Buffer	2.36	0.17	0.02	-0.01	0.05	-0.02
Deng Deng (CAM)	PA	7.26***	0.39	0.07	-0.04	0.32**	0.00
	Buffer	14.25***	0.56	0.11	-0.01	0.58***	-0.11
San Miguel de los Farallones (COL)	PA						
	Buffer	6.33**	0.36	0.33***	-0.18**	0.01	-0.04
Congolón, Piedra Parada y Coyocutena (HON)	PA	5.49**	0.33	0.50***	-0.37*	0.14	-0.06
	Buffer	1.50	0.12	0.43***	-0.07	0.14	-0.24
Montaña de Botaderos Carlos Escaleras Mejía (HON)	PA	7.45***	0.40	0.56*	0.17	0.52	0.86
	Buffer	4.11*	0.27	0.59***	-0.05	0.49*	0.20
Papikonda (IND)	PA	10.98***	0.49	0.14***	-0.12**	0.06	-0.01
	Buffer	4.56**	0.29	0.21***	-0.15**	-0.02	0.13*
Kyauk Pan Taung (MYA)	PA	17.12***	0.60	0.21*	-0.18	0.34*	0.55**
	Buffer	27.41***	0.71	0.27	-0.16	0.57*	1.71***
Bosques Nublados de Udima (PER)	PA	0.74	0.06	0.24*	0.11	0.03	-0.22
	Buffer	1.71	0.13	0.18*	0.19	0.04	-0.26
Mount Balatukan Range (PHI)	PA	6.88***	0.38	0.17*	-0.17	0.16	-0.08
	Buffer	9.71***	0.46	0.22**	-0.16	0.24**	-0.14

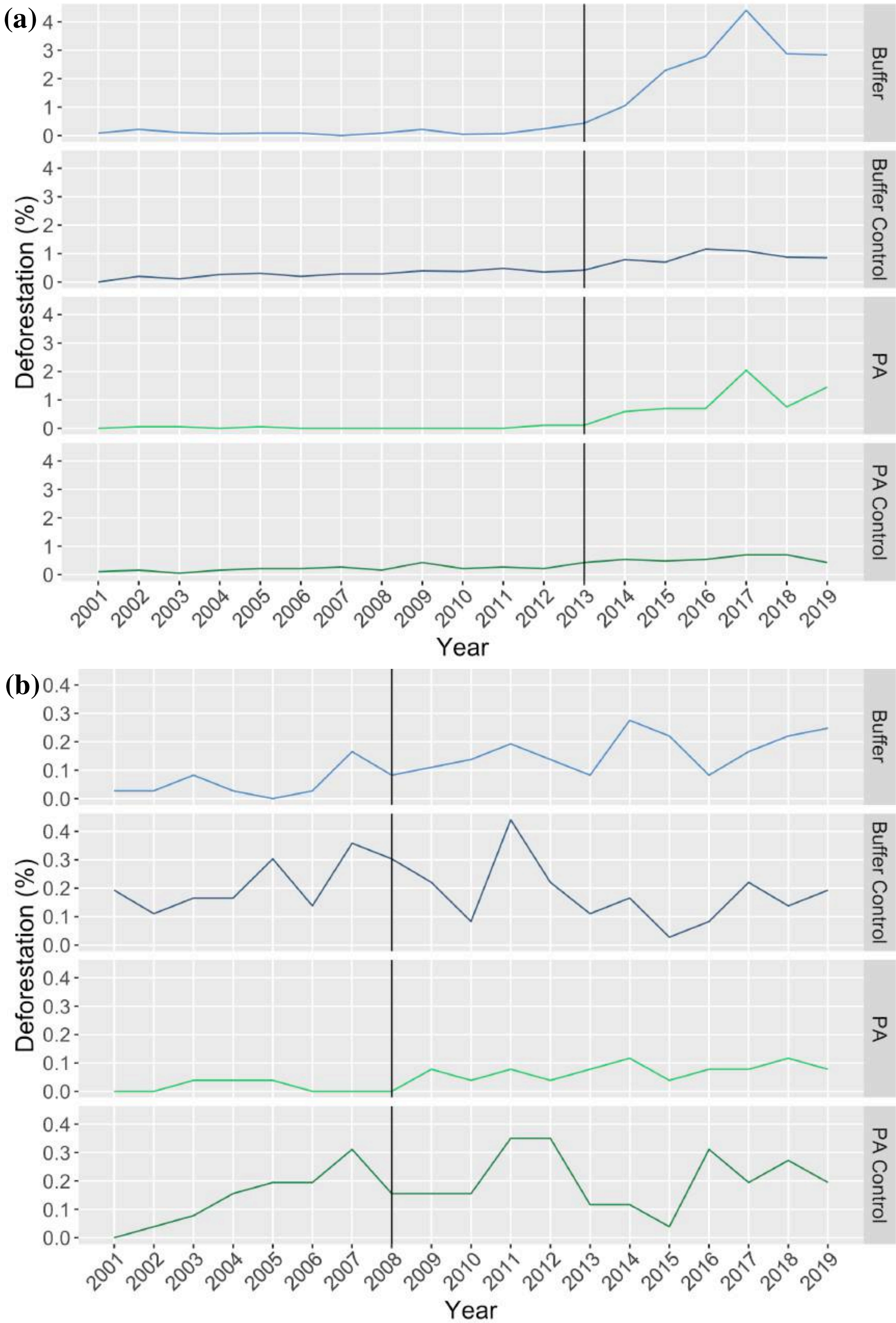

Figure 8. Deforestation across the study period for (a) Kyauk Pan Taun (MYA) and (b) Papikonda (IND) for the PA, buffer, and respective controls. The black vertical line denotes the year of PA designation.

4.4 Local spatial patterns and drivers of leakage

As Figure 9 shows, Kyauk Pan Taung (MYA) experienced a large increase in deforestation post-designation, within the PA, buffer zone, and beyond. The pattern of rates increasing from within the PA and peaking close to the edge of the buffer before declining is seen both before designation (Figure 9a), after designation (Figure 9b), and when comparing both to the matched control mean rate (Figure 9c). The major difference is that the post-designation deforestation curve is much steeper in gradient and greatly exceeds the control rates.

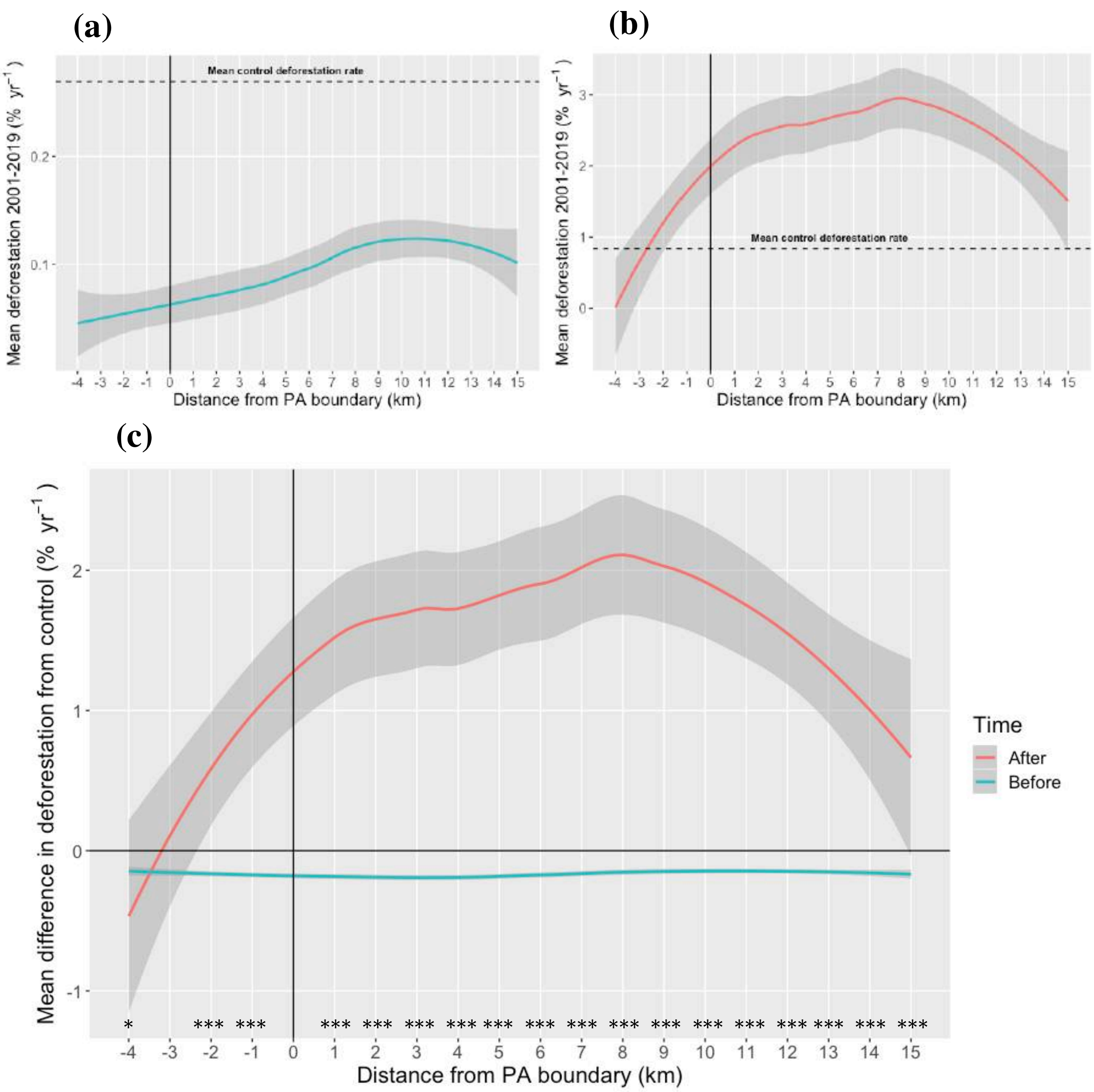

Figure 9. Kyauk Pan Taung (MYA) rates of deforestation (2001-2019) within 1km concentric rings from 4km within the PA to 15km outside of PA. Mean rate of deforestation before (a) and after (b) PA designation are shown with the mean difference from the respective control for both before and after (c). Lines were smoothed using a LOESS function, with the 95% confidence interval shown in grey. Significant difference (Mann-Whitney U test) between before and after for each distance is denoted using * (<0.05), ** (<0.01), and *** (<0.001).

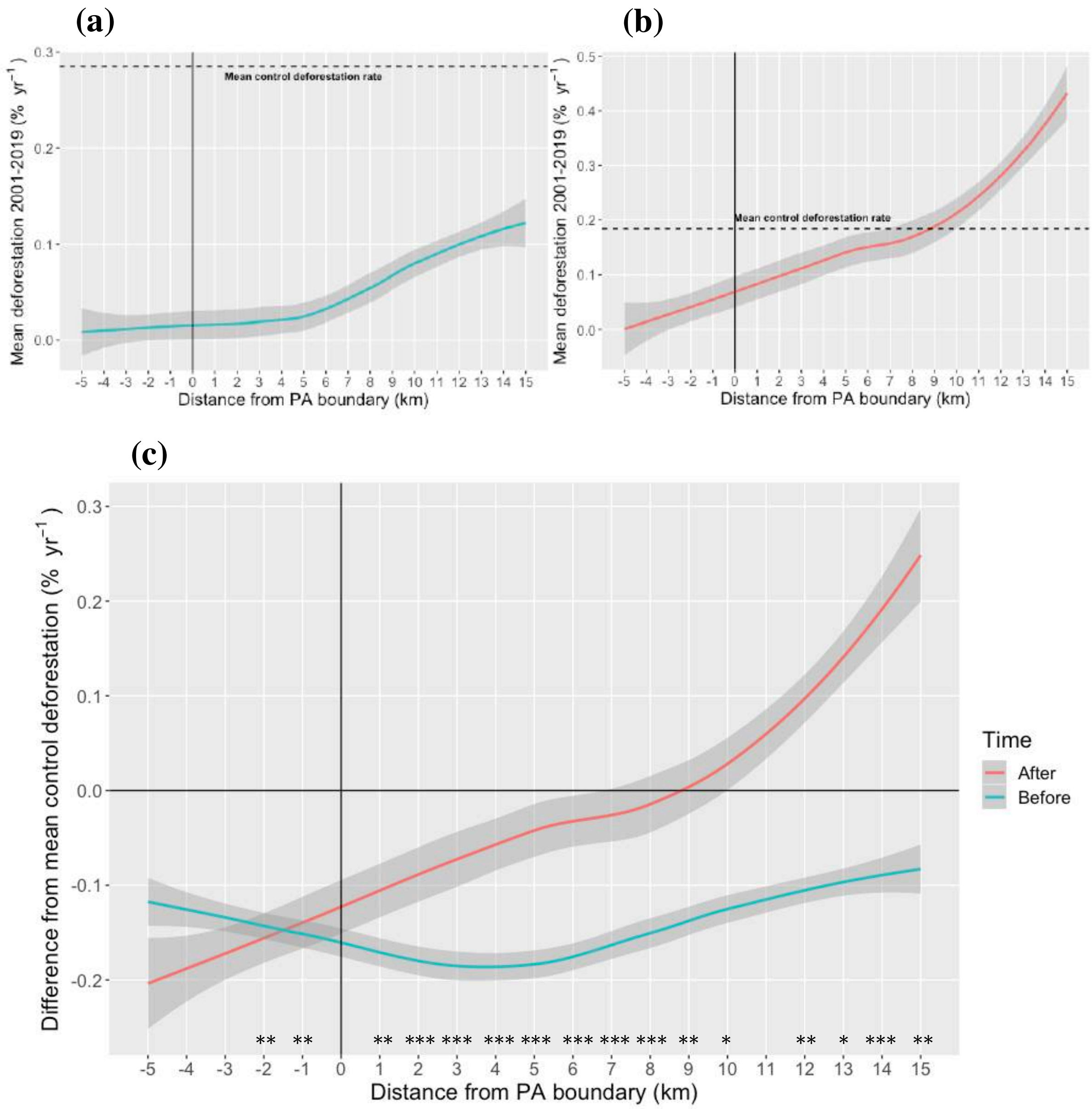

Figure 10. Papikonda (IND) rates of deforestation (2001-2019) within 1km concentric rings from 5km within the PA to 15km outside of PA. Mean rate of deforestation before (a) and after (b) PA designation are shown with the mean difference from the respective control for both before and after (c). Lines were smoothed using a LOESS function, with the 95% confidence interval shown in grey. Significant difference (Mann-Whitney U test) between before and after for each distance is denoted using * (<0.05), ** (<0.01), and *** (<0.001).

In contrast, Papikonda (IND) (Figure 10) has a gradual, almost linear, increase from within the PA to the edge of the buffer, remaining below the mean control deforestation rate both before and after designation. Rates are still higher post-designation and beyond the 10km buffer there is a sharp increase in deforestation.

As can be seen in Figure 11, the two PAs are very different in terms of accessibility and human settlement. Figure 11 also shows clearly how in Myanmar deforestation is associated with rivers, which resulted in the inclusion of rivers in this portion of the analysis. The logistic regression results in Table 8 show that before designation, deforestation in Kyauk Pan Taung (MYA) buffer zone was primarily linked with slope; however, after designation slope, elevation, and distance to river all had a significant negative relationship with deforestation probability. In Papikonda (IND) before designation distance to settlement, river, and slope were significantly negatively related to deforestation, post-designation was the same with the surprising addition of a significant positive relationship with elevation. The models generally performed well, all with significant likelihood ratio tests and low misclassification error. Significant coefficients contributed to the fit of the overall model (Residual deviance) with the exception of distance to settlement in Papikonda (IND) post-designation.

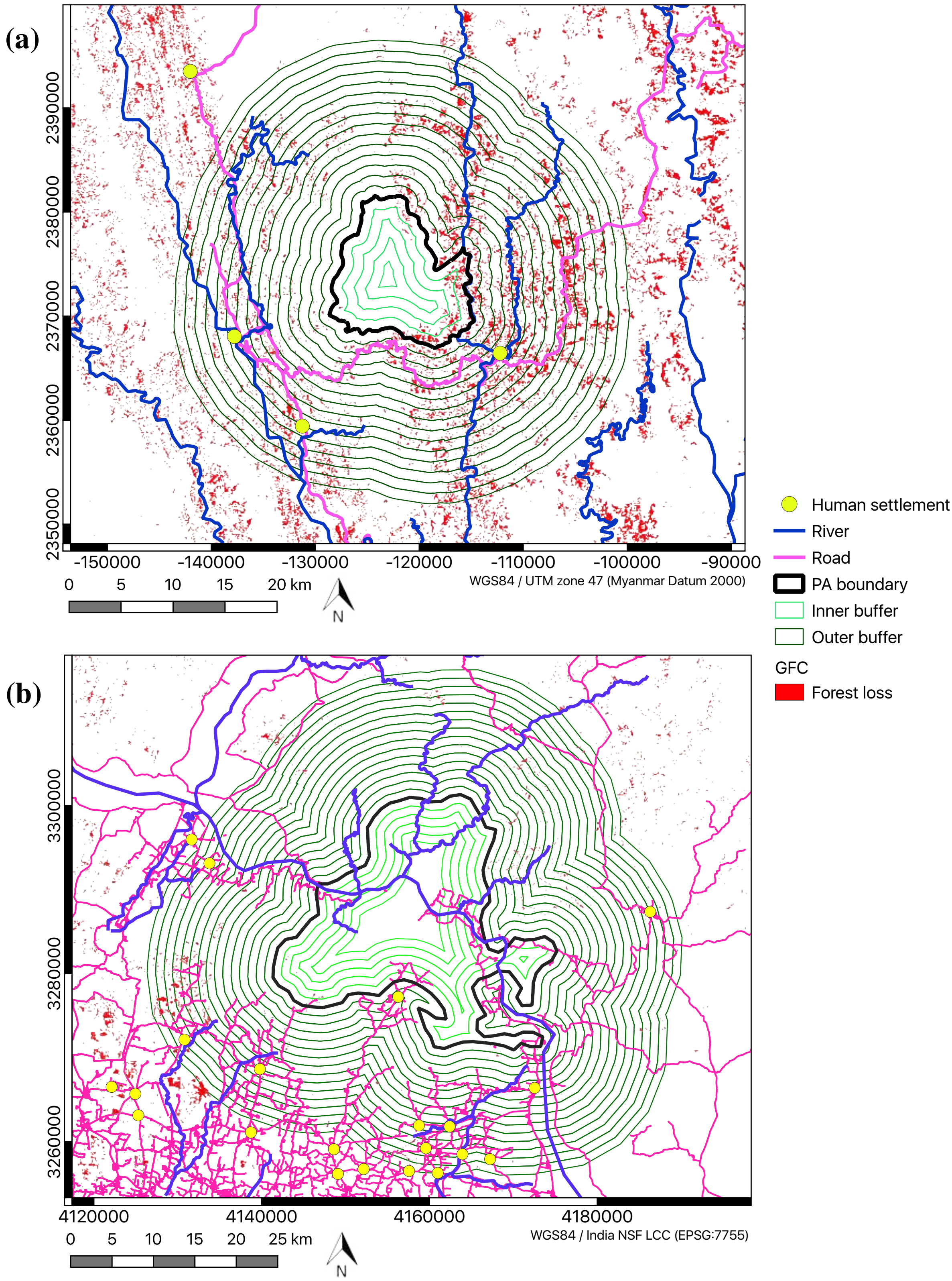

Figure 11. Maps showing the forest loss, transport infrastructure, and human settlements within the concentric rings of PA buffer used in the spatial analysis for (a) Kyauk Pan Taun (MYA) and (b) Papikonda (IND).

Table 8. Results of the logistic regression models for Kyauk Pan Taun (MYA) and Papikonda (IND) before and after PA designation. Coefficients of infrastructure accessibility were regressed against the presence or absence of deforestation. Model validity was assessed using the residual deviance of each coefficient and the misclassification error (see main text). Significant p-values (≤0.05) highlighted in bold.

			Intercept	**Distance to settlement**	**Distance to PA boundary**	**Distance to river**	**Distance to road**	**Elevation**	**Slope**	**Mis-classification error**	**Likelihood ratio test (p-value)**
Kyauk Pan Taun (MYA)	Before	Estimate	-3.194	3.759e-05	6.513e-05	-1.796e-04	-4.083e-05	-1.042e-03	-2.853e-02	0.012	**<0.001**
		Coefficient p-value	**<0.001**	0.523	0.264	0.071	0.546	0.418	**0.002**		
		Residual deviance p-value	N/A	0.170	**0.029**	**0.007**	0.795	0.078	**<0.001**		
	After	Estimate	-0.483	-1.491e-05	-1.336e-05	-2.542e-04	-3.013e-05	1.328e-03	-1.175e-02	0.162	**<0.001**
		Coefficient p-value	**0.001**	0.384	0.404	**<0.001**	0.121	**<0.001**	**<0.001**		
		Residual deviance p-value	N/A	**<0.001**	**0.012**	**<0.001**	**0.039**	**0.003**	**<0.001**		
Papikonda (IND)	Before	Estimate	-1.439	-3.487e-04	-7.383e-05	-3.188e-04	3.881e-04	2.718e-03	-8.148e-02	0.002	**0.002**
		Coefficient p-value	0.338	**0.027**	0.610	**0.034**	0.229	0.381	**0.019**		
		Residual deviance p-value	N/A	**0.045**	0.874	**0.032**	0.058	0.708	**0.003**		
	After	Estimate	-2.262	-1.101e-04	2.465e-05	-3.141e-04	4.972e-05	4.611e-03	-4.622e-02	0.017	**<0.001**
		Coefficient p-value	**<0.001**	**<0.001**	0.610	**<0.001**	0.441	**<0.001**	**<0.001**		
		Residual deviance p-value	N/A	0.936	0.526	**<0.001**	**0.012**	**<0.001**	**<0.001**		

5. Discussion

5.1 Matching

The use of matching to generate control samples was key to performing counterfactual analysis, however the balance of the variables was not perfect, as shown by the QQ plots in Appendix E; there are a number of potential reasons for this and clear opportunities for refinement in further work.

First of all, the GRUMP dataset of human settlements may lack the appropriate level of resolution for this analysis as discovered when investigation of the poor balance of the settlement distance covariate for the Mount Balatukan Range (PHI) PA, where on an island of 20 million inhabitants (Philippines Statistics Authority, 2015) there were only 32 settlements. Additionally, during the further spatial analysis of leakage in Kyauk Pan Taung (MYA) the small settlements in proximity to the PA and buffer were also not included in the GRUMP data. While these data may represent a good metric of access to major markets, for these fine-scale deforestation processes, it may be of more value and improve matching balance to include smaller settlements in the future.

The failure of San Miguel de los Farallones PA to achieve balance may be due to its small area (see Table 4). Joppa *et al.* (2008) specifically excluded PAs under 100km^2 as it had been found in some cases that these smaller parks are less effective (this is in some contention (Clark *et al.*, 2008; Ribas *et al.*, 2020) but there has been a number of supporting findings (Armenteras *et al.*, 2009; Geldmann *et al.*, 2015)); however this seems prematurely exclusionary and it could be argued that small PAs can be important ecological reservoirs in non-remote places (Geldmann *et al.*, 2015). That being said, it may be that matching becomes unviable when the homogeneity of landscape within small PAs results in poor quality of matches, which may have been the case in this study, considering the three smallest PAs all had issues achieving satisfactory balance.

An assumption with matching that must be considered is that balance in observed covariates is synonymous with balance in unobserved covariates (Schleicher *et al.*, 2020); as a result, the models should attempt to be as comprehensive as possible. However, the findings in Kyauk Pan Taung (MYA) and Papikonda (IND) that distance to river impacts deforestation and accessibility (Section 4.4) show that this was clearly not achieved in this analysis. The removal of covariates from Congolón, Piedra Parada y Coyocutena (HON) and Mount Balatukan Range (PHI) also jeopardise this requirement.

Additionally, the deforestation rates in the majority of control groups were greater than the respective PA/buffer area prior to (and in many after) designation. This could mean that the PAs represent remote regions with the least possible amount of human pressure in the study areas, or that there are additional significant confounding variables not included in the analysis; the distance to river finding suggests the latter may be more likely and it is recommended that a more comprehensive set of covariates is tested in further work. However, this study's use of BACI does offer a unique assessment of matching in the context of PA deforestation; the direct comparison of matched control samples to pre-treatment rates acts as a secondary test of validity. This is not possible in the majority of other studies that only cover post-designation time periods, leaving the assessment of matching quality to comparison of standardised means and QQ plots. As could be the case in this study, these measures may be insufficient, compromising the accuracy of the counterfactual and any conclusions drawn, especially when many PAs are offering marginal benefits (e.g. Joppa and Pfaff, 2011; Spracklen et al., 2015; Geldmann et al., 2019).

5.2 Deforestation Rates

In general, the rates found are comparable to the FAO's (2020b) findings for the time periods covered (see Table 9). However, a major trend in this study was an increase in annual deforestation when comparing before and after designation; contrasting with the FAO's (2020b) findings that regionally and globally, overall deforestation rates have slowed over the past two decades. Perhaps this indicates that the more inaccessible, intact forest ecosystems that PAs disproportionately represent (Heino *et al*., 2015), are experiencing increased pressure due the shrinking of the available forest stock and continued demand for agricultural expansion driving up the value of the forest resource (Armenteras *et al*., 2017; Jayathilake *et al*., 2020). Alternatively, this could demonstrates the risk in global or regional summaries such as the FAO's, which can mask local or ecosystem-specific issues, especially when deforestation is highly variable at the ecosystem, national, and subnational level (Hansen *et al*., 2013; Heino *et al*., 2015; Poor *et al*., 2019); for example, the increasing trends found in this study are supported by other similar fine scale research in tropical forest ecosystems (Austin *et al*., 2017; Geldmann *et al*., 2019). The extreme case of Montaña de Botaderos Carlos Escaleras Mejía (HON), where deforestation massively increased only in the PA after designation, is likely due to Honduras' extractivism following a military coup in 2009 and the granting of mining concessions within the PA (Serrano *et al*., 2016; Bebbington *et al*., 2018).

Table 9. Forest area change for the major global regions over the 3 decades since 1990, taken from FAO (2020b).

Region/subregion	Forest area annual change					
	1990–2000		2000–2010		2010–2020	
	1 000 ha/yr	%	1 000 ha/yr	%	1 000 ha/yr	%
Eastern and Southern Africa	-1 345	-0.40	-1 773	-0.55	-1 907	-0.62
Northern Africa	-182	-0.47	-127	-0.34	-168	-0.47
Western and Central Africa	-1 748	-0.50	-1 503	-0.45	-1 862	-0.59
Total Africa	**-3 275**	**-0.45**	**-3 403**	**-0.49**	**-3 938**	**-0.60**
East Asia	1 917	0.88	2 332	0.97	1 901	0.73
South and Southeast Asia	-1 843	-0.58	-262	-0.09	-941	-0.31
Western and Central Asia	129	0.26	285	0.55	213	0.39
Total Asia	**202**	**0.03**	**2 355**	**0.39**	**1 173**	**0.19**
Europe excl. Russian Federation	763	0.40	585	0.30	330	0.16
Total Europe	**795**	**0.08**	**1 171**	**0.12**	**348**	**0.03**
Caribbean	85	1.34	69	0.97	39	0.51
Central America	-218	-0.81	-211	-0.85	-130	-0.56
North America	-160	-0.02	327	0.05	-57	-0.01
Total North and Central America	**-293**	**-0.04**	**184**	**0.02**	**-148**	**-0.02**
Total Oceania	**-165**	**-0.09**	**-231**	**-0.13**	**423**	**0.23**
Total South America	**-5 102**	**-0.54**	**-5 249**	**-0.58**	**-2 597**	**-0.30**
WORLD	**-7 838**	**-0.19**	**-5 173**	**-0.13**	**-4 739**	**-0.12**

5.3 Protected Area Effectiveness

The DiD models indicate that the treatment of legal PA designation had no significant effect on deforestation rates within the PA area (other than in Kyauk Pan Taung (MYA) with significant increase in deforestation, covered below). Therefore, the reduced rates of deforestation when compared to the counterfactual matched control samples demonstrate that these PAs have *de facto* protection due to innate characteristics of their location (even though matching should control for these characteristics, see Section 5.1). This contrasts with other matching-based impact evaluations that have found that in general PAs do convey benefits in terms of avoided deforestation, although less than when determined through traditional or 'naïve' methods (Andam *et al*., 2008; Gaveau *et al*., 2009; Joppa and Pfaff, 2011; Ribas *et al*., 2020). This could be due to a lag in the effect of designation on deforestation, as Joppa and Pfaff (2011) and Fuller *et al*. (2019) specify in their studies that recently designated PAs should be avoided because of this inertia. Research with longer study periods are recommended to allow for differences in trends to be detected, suggested by Fuller *et al.* (2019) as a gap in the current literature.

The distributions of the residuals and poor explanatory power (r^2) of the DiD models are concerning. Perhaps this is due to the small length of the sample period or the non-linear, almost stochastic trends observed in deforestation rates. The linear nature of DiD may be unsuitable for studying deforestation trends, however, longer study periods could prove to be more successful, if less feasible.

5.4 Leakage

From the DiD models, only two PAs had buffer zones that were significantly elevated in comparison to the control and pre-treatment groups. This suggests that leakage is not a widespread phenomenon within newly designated PAs, a similar conclusion to other studies that assessed already established PAs (Andam *et al.*, 2008; Spracklen *et al.*, 2015; Lui and Coomes, 2016; Robalino *et al.*, 2017; Fuller *et al.*, 2019). Many of these papers found that blockage was a more common outcome than leakage but this was not found to be the case: The buffer zones often lost less forest than the control groups, which is due to the disparity in the pre-treatment groups as discussed above and not due to blockage, as would have been revealed by the DiD analysis.

The further spatial analysis of deforestation within the PA, buffer, and beyond the buffer revealed some key differences between Kyauk Pan Taung (MYA) and Papikonda (IND). The expected pattern of deforestation rates increasing the most across the PA boundary (Spracklen *et al.*, 2015) was not seen, perhaps because there were also increases within both PAs. The spatial distribution for Papikonda (IND) did not represent leakage, as deforestation rates were higher beyond the buffer, but far beyond the matched control group mean. This could represent an issue with the matching sampling strategy, where selecting samples from the relevant administrative division (state in this case) is not representative of the human pressure on the immediate landscape surrounding the PA; in the state of Andhra Pradesh (IND) Reddy *et al.* (2016) found that the area of Papikonda and immediate surrounds represent a deforestation hotspot, therefore comparing to the whole state, even using matching, could produce a control sample with much lower deforestation rates than relevant to the buffer zone, simulating a false leakage effect. This highlights the difficulty in broad-scale analysis of deforestation when the dynamics and drivers of spillovers can be so locally specific (Pfaff and Robalino, 2017). Another possibility is that the spatial extent of leakage is far beyond a 10km buffer, which is an arbitrary distance perpetuated in the literature to maintain comparability between studies (Joppa and Pfaff, 2010; Lui and Coomes, 2016; Poor *et al*., 2019); this is a major gap in the literature regarding conservation intervention spillovers and further work should investigate buffer zones of different sizes, for example for ecological viability, functional buffer zones depend on PA size (Alexandre *et al*., 2010).

In contrast to Papikonda (IND), the deforestation pattern observed in Kyauk Pan Taung (MYA) is much more representative of the expected leakage distribution, with a clear peak of deforestation within the buffer descending towards the mean control rate beyond. However, the elevation of deforestation seen within the PA and its significance as determined by the DiD analysis complicate the 'diagnosis' of leakage: If the PA also experiences an elevation of deforestation within its boundaries then the observed or implied displacement of pressure onto the buffer may just be an artefact of accessibility as the demarcation of the PA is failing to act as a deterrent. This is shown through the logistic regression in Section 4.4, where the post-designation deforestation rates have no relationship with the distance from the PA boundary and a strong negative relationship with slope, elevation, and distance to river; corresponding with the fact that the local rural communities are reliant on waterways for transport and trade, and that the PA is located on an isolated massif (Naing *et al*., 2017). The increase in deforestation rate post-designation could be attributed to the political and socioeconomic upheaval occurring in Myanmar over the past decade: The relaxation of the police state since 2010 has resulted in increased economic growth and liberalisation (Kraas *et al*., 2020), a potential driver of deforestation in the extractive economy of Myanmar (Prescott *et al*., 2017); especially in Chin State (where Kyauk Pan Taung is located) where the native peoples have historically experienced heavy persecution, poor infrastructure, and the nation's highest poverty rates (Hoffstaedter, 2014; Central Statistics Organisation and The World Bank, 2019; Nau, 2019).

The results of the logistic regression for Papikonda (IND) showed a similar lack of relationship between deforestation and the PA boundary, however distance to settlement was also a significant driver. This could be demonstrating that population density and urbanisation are greater drivers in Andhra Pradesh state, where a far larger and relatively wealthier population reside (see Table 10). These demographic and socioeconomic differences are important to highlight as they can affect deforestation and spillovers; for example Pfaff and Robalino (2017) argue that tourism and workforce mobility (facilitated by quality infrastructure) can be key in providing local people alternative modes of employment from extractive industries and therefore preventing leakage. These factors could be affecting the dynamics of the two PAs in question: Andhra Pradesh state has a thriving domestic tourism industry (Goodwin and Chaudhury, 2017), low poverty, and relatively high urban development and infrastructure with a mainly rural population; Chin state on the other hand has very high poverty, little infrastructure, and a tourism industry in its infancy (Kraas *et al.*, 2020). The disparity in the magnitude of deforestation rate between the two PAs could be a result of these characteristics. Additionally, prior remote sensing work has found that overall deforestation is increasing in Chin state and slowing in Andhra Pradesh (Krishna *et al.*, 2014; Wang and Myint, 2016).

Table 10. Demographic data for Chin State (MYA) and Andhra Pradesh State (IND), sourced from Myanmar Information Management Unit (2020) and Directorate of Economics & Statistics (2019).

	Population	**Population density (persons/km^2)**	**Rural population (%)**	**Population in poverty (%)**
Chin state (MYA)	478,801	13	79.2	73.3
Andhra Pradesh State (IND)	49,577,103	306	70.6	9.2

5.5 Further Work

As previously mentioned, refinement of the matching process offers great potential for increasing the conclusiveness and validity of the results, by increasing the confounding variables and the quality of data sources. Buffer zone size is another major gap in the understanding of leakage, although most other authors are content to perpetuate an arbitrary distance for the sake of comparability. As the datasets providing high resolution deforestation data grow in duration it will be crucial to continue to monitor PAs over longer periods of time to justify claims of effectiveness and reveal how dynamics change over time, particularly from designation onwards. The lack of definitive conclusions regarding leakage found in this study have revealed that broad-scale analysis may not be fit for purpose when investigating neighbourhood spillover dynamics; when individual cases are assessed in more detail, it becomes clear that these are complex issues with a range of possible drivers and additional confounding factors. As Adams *et al*. (2019) argues, the lack of consideration of additional geophysical, ecological, socioeconomic, and political factors, quantitatively or qualitatively, undermines the results of such studies. One specific aspect that has not been covered in this study is management efficacy and resources (Bruner *et al*., 2001), a factor almost certainly relevant when comparing the two PAs above. Therefore, site-specific study design using fine resolution, local-scale data is recommended.

On a more general note, deforestation and tropical PAs dominate the impact evaluation literature – this bias should be addressed and non-forest ecosystems must be considered. Additionally even with forest ecosystems, the use of forest conversion is relatively crude and can result in missing biodiversity losses, potentially leading to a 'half-empty forest' scenario, especially in PAs with 'sustainable' use (Redford and Feinsinger, 2001). This could be particularly relevant in more developed countries where land conversion is less likely but threats to ecosystem function still exist (Leverington *et al*., 2010).

6. Conclusions

This study has found that newly designated tropical forest PAs do not convey immediate significant benefits in the form of avoided deforestation and that deforestation rates have increased across the study period. Spillovers were not common and although leakage may be occurring in a small minority, complex local dynamics make identification uncertain, questioning the relevance of the coarse, broad analyses that have been previously used. The validity of the statistical analyses undermine the major conclusions drawn from the results, but have revealed some key areas for refinement in further work. PA impact evaluation is a complex field that combines ecology, economics, and politics, requiring better understanding and innovation in order to keep pace with the massive expansion of the global PA network.

References

Adams, V. M., Barnes, M. & Pressey, R. L. (2019) Shortfalls in Conservation Evidence: Moving from Ecological Effects of Interventions to Policy Evaluation. *One Earth*. 1(1), 62–75. Available from: https://doi.org/10.1016/j.oneear.2019.08.017.

Alexandre, B., Crouzeilles, R., Grelle, C. & Grelle, E. (2010) How Can We Estimate Buffer Zones of Protected Areas? A Proposal Using Biological Data. *Brazilian Journal of Nature Conservation*. 8(2), 165–170.

Alix-Garcia, J. & Gibbs, H. K. (2017) Forest conservation effects of Brazil's zero deforestation cattle agreements undermined by leakage. *Global Environmental Change*. 47, 201–217. Available from: http://www.sciencedirect.com/science/article/pii/S095937801730170X.

Amin, A., Choumert-Nkolo, J., Combes, J.-L., Combes Motel, P., Kéré, E. N., Ongono-Olinga, J.-G. & Schwartz, S. (2019) Neighborhood effects in the Brazilian Amazônia: Protected areas and deforestation. *Journal of Environmental Economics and Management*. 93, 272–288. Available from: http://www.sciencedirect.com/science/article/pii/S0095069616303163.

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*. 105(42), 16089–16094. Available from: http://www.pnas.org/content/105/42/16089.abstract.

Anderson, C. M., Asner, G. P., Llactayo, W. & Lambin, E. F. (2018) Overlapping land allocations reduce deforestation in Peru. *Land Use Policy*. 79, 174–178. Available from: http://www.sciencedirect.com/science/article/pii/S0264837717311109.

Angrist, J. & Pischke, J.-S. (2008) 'Differences-in-differences: Pre and Post, Treatment and Control', in *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. 169. Available from: https://www.researchgate.net/publication/51992844_Mostly_Harmless_Econometrics_An_Empiricist%27s_Companion.

Armenteras, D., Espelta, J. M., Rodríguez, N. & Retana, J. (2017) Deforestation dynamics and drivers in different forest types in Latin America: Three decades of studies (1980–2010). *Global Environmental Change*. 46, 139–147. Available from: http://www.sciencedirect.com/science/article/pii/S0959378016304745.

Armenteras, D., Rodríguez, N. & Retana, J. (2009) Are conservation strategies effective in avoiding the deforestation of the Colombian Guyana Shield? *Biological Conservation*. 142(7), 1411–1419. Available from: http://www.sciencedirect.com/science/article/pii/S0006320709000767.

Sánchez-Azofeifa, G., Daily, G. C., Pfaff, A. S. P. & Busch, C. (2003) Integrity and isolation of Costa Rica's national parks and biological reserves: examining the dynamics of land-cover change. *Biological Conservation*. 109(1), 123–135. Available from: http://www.sciencedirect.com/science/article/pii/S0006320702001453.

Aung Moe, U. (2016) 'Geospatial Data Infrastructure in Myanmar'. *FIG Asia Pacific Capacity Development Network: Regional Capacity Network Meeting*. Kuala Lumpur, Malaysia, 17-19 October. Available from: fig.net/resources/proceedings/2016/2016_10_AP_CDN/1.2_Myanmar.pdf

Aukland, L., Costa, P. M. & Brown, S. (2003) A conceptual framework and its application for addressing leakage: the case of avoided deforestation. *Climate Policy*. 3(2), 123–136. Available from: http://www.sciencedirect.com/science/article/pii/S1469306202000657.

Austin, K. G., González-Roglich, M., Schaffer-Smith, D., Schwantes, A. M. & Swenson, J. J. (2017) Trends in size of tropical deforestation events signal increasing dominance of industrial-scale drivers. *Environmental Research Letters*. 12(5). Available from: http://dx.doi.org/10.1088/1748-9326/aa6a88.

Balk, D. L., Deichmann, U., Yetman, G., Pozzi, F., Hay, S. I. & Nelson, A. (2006) Determining Global Population Distribution: Methods, Applications and Data. *Advances in Parasitology*. 62, 119–156. Available from: https://doi.org/10.1016/S0065-308X(05)62004-0.

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(1), 58–54.

Bebbington, A., Fash, B. & Rogan, J. (2018) Socio-environmental Conflict, Political Settlements, and Mining Governance: A Cross-Border Comparison, El Salvador and Honduras. *Latin American Perspectives*. 46(2), 84–106. Available from: https://doi.org/10.1177/0094582X18813567.

Bennett, G. & Mulongoy, K. J. (2006) *Review of experience with ecological networks, corridors and buffer zones*. Secretariat of the Convention on Biological Diversity: Montreal, Technical Series No. 23. 7. Available from:

https://www.cbd.int/doc/publications/cbd-ts-23.pdf

Bingham, H. C., Juffe Bignoli, D., Lewis, E., MacSharry, B., Burgess, N. D., Visconti, P., Deguignet, M., Misrachi, M., Walpole, M., Stewart, J. L., Brooks, T. M. & Kingston, N. (2019) Sixty years of tracking conservation progress using the World Database on Protected Areas. *Nature Ecology & Evolution*. 3(5), 737–743. Available from: https://doi.org/10.1038/s41559-019-0869-3.

Bode, M., Tulloch, A. I. T., Mills, M., Venter, O. & Ando, A. W. (2015) A conservation planning approach to mitigate the impacts of leakage from protected area networks. *Conservation Biology*. 29(3), 765–774.

Brown, P., Cabarle, B. & Livernash, R. (1997) *Carbon counts: estimating climate change mitigation in forestry projects*. Washington D.C.: World Resources Institute

Bruner, A. G., Gullison, R. E., Rice, R. E. & da Fonseca, G. A. B. (2001) Effectiveness of Parks in Protecting Tropical Biodiversity. *Science*. 291(5501), 125–128. Available from: http://science.sciencemag.org/content/291/5501/125.abstract.

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M. & Palmer, T. M. (2015) Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*. 1(5), e1400253. Available from: http://advances.sciencemag.org/content/1/5/e1400253.abstract.

Center for International Earth Science Information Network (Columbia University), Institute for Demographic Research (City University New York), International Food Policy Research Institute, The World Bank, and Centro Internacional de Agricultura Tropical (2017) *Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Settlement Points, Revision 01*. NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY. Available from: https://doi.org/10.7927/H4BC3WG1 [Accessed June 2020]

Central Statistics Organisation and The World Bank (2019) *The Myanmar Living Conditions Survey Poverty Report*. The World Bank and Myanmar Ministry of Planning and Finance: Nay Pyi Taw, Myanmar. Available from: https://www.worldbank.org/en/country/myanmar/publication/poverty-report-myanmar-living-conditions-survey-2017

Chape, S., Harrison, J., Spalding, M. & Lysenko, I. (2005) Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 360(1454), 443–455. Available from: https://doi.org/10.1098/rstb.2004.1592.

Coad, L., Watson, J. E. M., Geldmann, J., Burgess, N. D., Leverington, F., Hockings, M., Knights, K. & Di Marco, M. (2019) Widespread shortfalls in protected area resourcing undermine efforts to conserve biodiversity. *Frontiers in Ecology and the Environment*. 17(5), 259–264. Available from: https://doi.org/10.1002/fee.2042.

Convention on Biological Diversity (2010) *Decision X/2: Strategic Plan for Biodiversity 2011-2020*. Convention on Biological Diversity: Nagoya, Japan.

Congedo, L. (2016) *Semi-Automatic Classification Plugin v6.2.6*. QGIS Python Plugins Repository. Available from: https://plugins.qgis.org/plugins/SemiAutomaticClassificationPlugin/

Crutzen, P. J. (2002) Geology of mankind. *Nature*. 415(6867), 23. Available from: https://doi.org/10.1038/415023a.

Curran, L. M., Trigg, S. N., McDonald, A. K., Astiani, D., Hardiono, Y. M., Siregar, P., Caniago, I. & Kasischke, E. (2004) Lowland Forest Loss in Protected Areas of Indonesian Borneo. *Science*. 303(5660), 1000–1003. Available from: http://science.sciencemag.org/content/303/5660/1000.abstract.

DeFries, R., Hansen, A., Newton, A. C. & Hansen, M. C. (2005) Increasing isolation of protected areas in tropical forests over the past twenty years. *Ecological Applications*. 15(1), 19–26. Available from: https://doi.org/10.1890/03-5258.

Delacote, P., Robinson, E. J. Z. & Roussel, S. (2016) Deforestation, leakage and avoided deforestation policies: A spatial analysis. *Resource and Energy Economics*. 45, 192–210. Available from: http://www.sciencedirect.com/science/article/pii/S0928765516301737.

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(3), 932–945. Available from: http://www.jstor.org/stable/43554804.

Directorate of Economics & Statistics (2019) *Statistical Abstract Andhra Pradesh 2019*. Government of Andhra Pradesh: Gollapudi, India. Available from https://desap.in/MainPage.do?mode=menuBind&tabname=publications

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B. & Collen, B. (2014) Defaunation in the Anthropocene. *Science*. 345(6195), 401–406. Available from: http://science.sciencemag.org/content/345/6195/401.abstract.

Dudley, N. (ed.) (2008) *Guidelines for Applying Protected Area Management Categories*. IUCN, Gland, Switzerland.

Ewers, R. M. & Rodrigues, A. S. L. (2008) Estimates of reserve effectiveness are confounded by leakage. *Trends in Ecology & Evolution*. 23(3), 113–116. Available from: http://www.sciencedirect.com/science/article/pii/S0169534708000402.

Fajardo, S. (2001) 'Surveying and Mapping in the Philippines'. *20th International Cartographic Conference*, Beijing, 6-10 August. Available from: https://dev.icaci.org/files/documents/national_reports/1999-2003/Philippines.pdf

FAO (2020a) *Forest Resources Assessment working paper 188: Terms and definitions*. Food and Agriculture Organisation of the United Nations: Rome, Italy. p4. Available from www.fao.org/3/I8661EN/i8661en.pdf

FAO (2020b) *Global Forest Resources Assessment 2020: Main Report*. Food and Agriculture Organisation of the United Nations: Rome, Italy. p17. Available from: http://www.fao.org/3/ca9825en/CA9825EN.pdf

Ferraro, P. J. (2009) Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation*. 2009(122), 75–84. Available from: https://doi.org/10.1002/ev.297.

Ferraro, P. J. & Pattanayak, S. K. (2006) Money for Nothing? A Call for Empirical Evaluation of Biodiversity Conservation Investments. *PLOS Biology*. 4(4), e105. Available from: https://doi.org/10.1371/journal.pbio.0040105.

Ferraro, P. J. & Pressey, R. L. (2015) Measuring the difference made by conservation initiatives: protected areas and their environmental and social impacts. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 370(1681), 20140270. Available from: https://doi.org/10.1098/rstb.2014.0270.

Fox, J. & Weisberg, S. (2019). *An {R} Companion to Applied Regression*, 3rd Edition. Thousand Oaks CA: Sage. Available from: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fuller, C., Ondei, S., Brook, B. W. & Buettel, J. C. (2019) First, do no harm: A systematic review of deforestation spillovers from protected areas. *Global Ecology and Conservation*. 18, e00591. Available from: http://www.sciencedirect.com/science/article/pii/S2351989419301143.

Gan, J. & McCarl, B. A. (2007) Measuring transnational leakage of forest conservation. *Ecological Economics*. 64(2), 423–432. Available from:

http://www.sciencedirect.com/science/article/pii/S0921800907001668.

Gardner, T. A., Barlow, J., Sodhi, N. S. & Peres, C. A. (2010) A multi-region assessment of tropical forest biodiversity in a human-modified world. *Biological Conservation*. 143(10), 2293–2300. Available from: http://www.sciencedirect.com/science/article/pii/S0006320710002533.

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(11), 2165–2175. Available from: https://doi.org/10.1111/j.1365-2699.2009.02147.x.

Geldmann, J., Barnes, M., Coad, L., Craigie, I. D., Hockings, M. & Burgess, N. D. (2013) Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biological Conservation*. 161, 230–238. Available from: http://www.sciencedirect.com/science/article/pii/S0006320713000670.

Geldmann, J., Coad, L., Barnes, M., Craigie, I. D., Hockings, M., Knights, K., Leverington, F., Cuadros, I. C., Zamora, C., Woodley, S. & Burgess, N. D. (2015) Changes in protected area management effectiveness over time: A global analysis. *Biological Conservation*. 191, 692–699. Available from: http://www.sciencedirect.com/science/article/pii/S0006320715300793.

Geldmann, J., Manica, A., Burgess, N. D., Coad, L. & Balmford, A. (2019) A global-level assessment of the effectiveness of protected areas at resisting anthropogenic pressures. *Proceedings of the National Academy of Sciences*. 116(46), 23209–23215. Available from: http://www.pnas.org/content/116/46/23209.abstract.

Ghosh, J. K. & Dubey, A. (2008) India's new map policy – Utility of civil users. *Current Science*. 94(3), 332–337. Available from: http://www.jstor.org/stable/24100340.

Giam, X. (2017) Global biodiversity loss from tropical deforestation. *Proceedings of the National Academy of Sciences*. 114(23), 5775–5777. Available from: http://www.pnas.org/content/114/23/5775.abstract.

Goodwin, R. D., Chaudhury, S. K. (2017) Sustainable promotional strategies of eco-tourism and its development in the sun rise state of Andhra Pradesh. *Asia Pacific Journal of Research*. 1(56), 195-202. Available from: https://www.researchgate.net/profile/Suman_Chaudhury/publication/340809644_SUSTAINABLE_PROMOTIONAL_STRATEGIES_OF_ECO-TOURISM_AND_ITS_DEVELOPMENT_IN_THE_SUN_RISE_STATE_OF_ANDH

RA_PRADESH/links/5e9eccfc4585150839ef666d/SUSTAINABLE-PROMOTIONAL-STRATEGIES-OF-ECO-TOURISM-AND-ITS-DEVELOPMENT-IN-THE-SUN-RISE-STATE-OF-ANDHRA-PRADESH.pdf

Google (2020a) Map data: Google, Maxar Technologies - Myanmar 21.30, 93.00°. Accessed using QGIS XYZ Tile importer, URL: https://mt1.google.com/vt/lyrs=y&x={x}&y={y}&z={z}

Google (2020b) Map data: Google, Maxar Technologies - India 17.40, 81.50°. Accessed using QGIS XYZ Tile importer, URL: https://mt1.google.com/vt/lyrs=y&x={x}&y={y}&z={z}

Hansen, M. C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. & Townshend, J. R. G. (2013) High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*. 342(6160), 850–853. Available from: http://earthenginepartners.appspot.com/science-2013-global-forest.

Heino, M., Kummu, M., Makkonen, M., Mulligan, M., Verburg, P. H., Jalava, M. & Räsänen, T. A. (2015) Forest Loss in Protected Areas and Intact Forest Landscapes: A Global Analysis. *PLOS ONE*. 10(10), e0138918. Available from: https://doi.org/10.1371/journal.pone.0138918.

Henders, S. & Ostwald, M. (2014) Accounting methods for international land-related leakage and distant deforestation drivers. *Ecological Economics*. 99, 21–28. Available from: http://www.sciencedirect.com/science/article/pii/S0921800914000093.

Herrera, D., Pfaff, A. & Robalino, J. (2019) Impacts of protected areas vary with the level of government: Comparing avoided deforestation across agencies in the Brazilian Amazon. *Proceedings of the National Academy of Sciences*. 116(30), 14916–14925. Available from: http://www.pnas.org/content/116/30/14916.abstract.

Ho, D., Imai, K., King, G., Stuart, E. (2007) Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*. 15(3), 199-236. Available from: http://gking. harvard.edu/files/abs/matchp-abs.shtml.

Ho, D., Imai, K., King, G., Stuart, E. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1-28. Available from: http://www.jstatsoft.org/v42/i08/

Hoffstaedter, G. (2014) Place-making: Chin refugees, citizenship and the state in Malaysia. *Citizenship Studies*. 18(8), 871–884. Available from:

https://doi.org/10.1080/13621025.2014.964549.

Hosmer, D., Lemeshow, S. & Sturdivant, R. (2013a) ‘Introduction to the Logistic Regression’, in *Applied Logistic Regression*. 3rd edition Hobo: John Wiley & Sons, Ltd. p. 30.

Hosmer, D., Lemeshow, S. & Sturdivant, R. (2013b) ‘Assessing the Fit of the Model’, in *Applied Logistic Regression*. 3rd edition Hoboken: John Wiley & Sons, Ltd. p. 155.

IUCN and UNEP-WCMC (2020a), The World Database on Protected Areas (WDPA). Cambridge, UK: UNEP-WCMC. Available at: www.protectedplanet.net.

Jayathilake, H. M., Prescott, G. W., Carrasco, L. R., Rao, M. & Symes, W. S. (2020) Drivers of deforestation and degradation for 28 tropical conservation landscapes. *Ambio*. Available from: https://doi.org/10.1007/s13280-020-01325-9.

Jenkins, C. N. & Joppa, L. (2009) Expansion of the global terrestrial protected area system. *Biological Conservation*. 142(10), 2166–2174. Available from: http://www.sciencedirect.com/science/article/pii/S0006320709001980.

Joppa, L. N., Loarie, S. R. & Pimm, S. L. (2008) On the protection of “protected areas”. *Proceedings of the National Academy of Sciences*. 105(18), 6673 LP – 6678. Available from: http://www.pnas.org/content/105/18/6673.abstract.

Joppa, L. N. & Pfaff, A. (2011) Global protected area impacts. *Proceedings of the Royal Society B: Biological Sciences*. 278(1712), 1633–1638. Available from: https://doi.org/10.1098/rspb.2010.1713.

Joppa, L. & Pfaff, A. (2010) Reassessing the forest impacts of protection: The challenge of nonrandom location and a corrective method. *Annals of the New York Academy of Sciences*. 1185(1), 135–149. Available from: https://doi.org/10.1111/j.1749-6632.2009.05162.x.

Jurgiel, B. (2020) *Point sampling tool v0.5.3*. QGIS Python Plugins Repository. Available from: https://plugins.qgis.org/plugins/pointsamplingtool/

Kande, L., Kamguia, J., Yap, L. and Foyang, L. (2016) Reference Ellipsoid Parameters of Cameroon from GPS Data. *International Journal of Geosciences*, **7**, 1049-1056. Available from: 10.4236/ijg.2016.78079.

Kraas, F., Myint, Z. & Hausler, N. (2020) ‘Tourism development in Myanmar: Dynamics, policies, and challenges’, in *Tourism and Development in Southeast Asia*. 1st edition Abingdon: Routledge. p.

Krishna, P. H., Saranya, K. R. L., Reddy, C. S., Jha, C. S. & Dadhwal, V. K. (2014)

Assessment and monitoring of deforestation from 1930 to 2011 in Andhra Pradesh, India using remote sensing and collateral data. *Current Science*. 107(5), 867–875. Available from: https://www.jstor.org/stable/24105592.

Laurance, W., Carolina Useche, D., Rendeiro, J., Kalka, M., Bradshaw, C., Sloan, S. P., Laurance, S. G., Campbell, M., Abernethy, K., Alvarez, P., Arroyo-Rodriguez, V., Ashton, P., Benítez-Malvido, J., Blom, A., Bobo, K. S., Cannon, C. H., Cao, M., Carroll, R., Chapman, C., Coates, R. *et al*. (2012) Averting biodiversity collapse in tropical forest protected areas. *Nature*. 489(7415), 290–294. Available from: https://doi.org/10.1038/nature11318.

Lechner, M. (2011) The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends in Econometrics*. 4(3), 165–224.

Leverington, F., Costa, K. L., Pavese, H., Lisle, A. & Hockings, M. (2010) A Global Analysis of Protected Area Management Effectiveness. *Environmental Management*. 46(5), 685–698. Available from: https://doi.org/10.1007/s00267-010-9564-5.

Lim, C. L., Prescott, G. W., De Alban, J. D. T., Ziegler, A. D. & Webb, E. L. (2017) Untangling the proximate causes and underlying drivers of deforestation and forest degradation in Myanmar. *Conservation Biology*. 31(6), 1362–1372. Available from: https://doi.org/10.1111/cobi.12984.

Liu, Y., Fu, B., Wang, S. & Zhao, W. (2018) Global ecological regionalization: from biogeography to ecosystem services. *Current Opinion in Environmental Sustainability*. 33, 1–8. Available from: http://www.sciencedirect.com/science/article/pii/S187734351730218X.

Longley, P., Goodchild, M., Maguire, D. & Rhind, D. (2015) 'Chapter 4: Georeferencing', in *Geographic Information Science And Systems*. 4th edition Hoboken: Wiley. p. 92.

Lui, G. V & Coomes, D. A. (2016) Tropical nature reserves are losing their buffer zones, but leakage is not to blame. *Environmental Research*. 147, 580–589. Available from: http://www.sciencedirect.com/science/article/pii/S0013935115301365.

Macdonald, E. A., Collins, M., Johnson, P. J., Clayton, L. M., Malhi, Y., Fisher, J. B., Milner-Gulland, E. J. & Macdonald, D. W. (2011) Wildlife conservation and reduced emissions from deforestation in a case study of Nantu National Park, Sulawesi: 1. The effectiveness of forest protection—many measures, one goal. *Environmental Science & Policy*. 14(6), 697–708. Available from: http://www.sciencedirect.com/science/article/pii/S1462901111000232.

Martínez W.A., Sánchez L. (2009) Realization of the SIRGAS Reference Frame in Colombia. In: Drewes H. (eds) *Geodetic Reference Frames. International Association of Geodesy Symposia*, vol 134. Springer, Berlin, Heidelberg

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Policy Responses: Chapter 5. Biodiversity*. Island Press: Washington DC. p122. Available from: https://www.millenniumassessment.org/en/Responses.html

Minghini, M. & Frassinelli, F. (2019) OpenStreetMap history for intrinsic quality assessment: Is OSM up-to-date? *Open Geospatial Data, Software and Standards*. 4(1), 9. Available from: https://doi.org/10.1186/s40965-019-0067-x.

Mitchard, E. T. A. (2018) The tropical forest carbon cycle and climate change. *Nature*. 559(7715), 527–534. Available from: https://doi.org/10.1038/s41586-018-0300-2.

Moilanen, A. & Laitila, J. (2016) Indirect leakage leads to a failure of avoided loss biodiversity offsetting. *Journal of Applied Ecology*. 53(1), 106–111. Available from: https://doi.org/10.1111/1365-2664.12565.

Mugnier, C. J. (2006) Grids & Datums Republic of Peru. *Photogrammetric Engineering and Remote Sensing*. 495–496. Available from: https://www.academia.edu/download/37583275/Peru.pdf.

Myanmar Information Management Unit (2020) *Chin Baseline Data*. Myanmar Information Management Unit, United Nations Resident and Humanitarian Coordinator: Yangon, Myanmar. Available from: https://themimu.info/states_regions/chin

Naing, T. Z., Lin, N., Platt, S., Platt, K., Khaing, L. L., Yu, T. T., Myo, K., Soe, M. M., Nwe, S. S. & Tizard, R. (2017) Birds of the Chindwin River and adjacent areas of the Chin and Naga Hills in western Myanmar. *Forktail*. 33, 64–76. Available from: https://www.researchgate.net/publication/326987999_Birds_of_the_Chindwin_River_and_adjacent_areas_of_the_Chin_and_Naga_Hills_in_western_Myanmar.

Nasiri, A., Abbaspour, R. A., Chehreghan, A. & Arsanjani, J. J. (2018) Improving the quality of citizen contributed geodata through their historical contributions: The case of the road network in OpenStreetMap. *ISPRS International Journal of Geo-Information*. 7(7), 253. Available from: https://doi.org/10.3390/ijgi7070253.

National Administrative Department of Statistics (2018) DANE Geoportal: National Geostatistical Framework [Online], Bogota, Colombia: DANE. Available from https://geoportal.dane.gov.co/servicios/descarga-y-metadatos/descarga-mgn-marco-geoestadistico-nacional/ [Accessed June 2020]

Nau, S.Z. 2019. The Areas Needed to Strengthen Land Governance for Sustainable Rural Development in Myanmar. *Sustainable Development Goals: Land Governance for Equivalent Living and Working Conditions in Urban and Rural Areas*, 16-18 May 2019, Qingdao, China. Available from: https://espace.curtin.edu.au/handle/20.500.11937/76487

Negret, P. J., Di-Marco, M., 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*. n/a(n/a), . Available from: 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(8), e22722. Available from: https://doi.org/10.1371/journal.pone.0022722.

Oldekop, J. A., Sims, K. R. E., Karna, B. K., Whittingham, M. J. & Agrawal, A. (2019) Reductions in deforestation and poverty from decentralized forest management in Nepal. *Nature Sustainability*. 2(5), 421–428. Available from: https://doi.org/10.1038/s41893-019-0277-3.

Oliveira, P. J. C., Asner, G. P., Knapp, D. E., Almeyda, A., Galván-Gildemeister, R., Keene, S., Raybin, R. F. & Smith, R. C. (2007) Land-Use Allocation Protects the Peruvian Amazon. *Science*. 317(5842), 1233–1236. Available from: http://science.sciencemag.org/content/317/5842/1233.abstract.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. & Kassem, K. R. (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*. 51(11), 933–938. Available from: https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO.

Peng, C.-Y. J., Lee, K. L. & Ingersoll, G. M. (2002) An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*. 96(1), 3–14. Available from: https://doi.org/10.1080/00220670209598786.

Pfaff, A. & Robalino, J. (2017) Spillovers from Conservation Programs. *Annual Review of Resource Economics*. 9(1), 299–315. Available from: https://doi.org/10.1146/annurev-resource-100516-053543.

Philippines Statistics Authority (2015) *Census of Population and Housing: Population of Legislative Districts by Region, Province, and Selected Highly Urbanized/Component City*. Philippines Statistics Authority: Quezon City, Philippines. Available from https://psa.gov.ph/population-and-housing/node/60002

Poor, E. E., Frimpong, E., Imron, M. A. & Kelly, M. J. (2019) Protected area effectiveness in a sea of palm oil: A Sumatran case study. *Biological Conservation*. 234, 123–130. Available from: http://www.sciencedirect.com/science/article/pii/S0006320718313399.

Potapov, P. V, Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Altstatt, A., Mane, L. & Justice, C. O. (2012) Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sensing of Environment*. 122, 106–116. Available from: http://www.sciencedirect.com/science/article/pii/S0034425712000430.

Prem, M., Saavedra, S. & Vargas, J. F. (2020) End-of-conflict deforestation: Evidence from Colombia's peace agreement. *World Development*. 129, 104852. Available from: http://www.sciencedirect.com/science/article/pii/S0305750X19305017.

Prescott, G. W., Sutherland, W. J., Aguirre, D., Baird, M., Bowman, V., Brunner, J., Connette, G. M., Cosier, M., Dapice, D., De Alban, J. D. T., Diment, A., Fogerite, J., Fox, J., Hlaing, W., Htun, S., Hurd, J., LaJeunesse Connette, K., Lasmana, F., Lim, C. L., Lynam, A. *et al*. (2017) Political transition and emergent forest-conservation issues in Myanmar. *Conservation Biology*. 31(6), 1257–1270. Available from: https://doi.org/10.1111/cobi.13021.

QGIS Development Team (2020). *QGIS Geographic Information System Version 3.4*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Qiu, J., Game, E. T., Tallis, H., Olander, L. P., Glew, L., Kagan, J. S., Kalies, E. L., Michanowicz, D., Phelan, J., Polasky, S., Reed, J., Sills, E. O., Urban, D. & Weaver, S. K. (2018) Evidence-Based Causal Chains for Linking Health, Development, and Conservation Actions. *BioScience*. 68(3), 182–193. Available from: https://doi.org/10.1093/biosci/bix167.

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

Rasolofoson, R. A., Ferraro, P. J., Jenkins, C. N. & Jones, J. P. G. (2015) Effectiveness of Community Forest Management at reducing deforestation in Madagascar. *Biological Conservation*. 184, 271–277. Available from:

http://www.sciencedirect.com/science/article/pii/S0006320715000488.

Reddy, C. S., Manaswini, G., Satish, K. V, Singh, S., Jha, C. S. & Dadhwal, V. K. (2016) Conservation priorities of forest ecosystems: Evaluation of deforestation and degradation hotspots using geospatial techniques. *Ecological Engineering*. 91, 333–342. Available from: http://www.sciencedirect.com/science/article/pii/S0925857416301562.

Redford, Kent & Feinsinger, P. (2001) 'The half-empty forest: Sustainable use and the ecology of interactions', in Reynolds, J., G. Mace, KH Redford & J. Robinson (eds.) *Conservation of Exploited Species*. Cambridge: Cambridge University Press. pp. 370–399.

Renwick, A. R., Bode, M. & Venter, O. (2015) Reserves in Context: Planning for Leakage from Protected Areas. *PLOS one*. 10(6), e0129441–e0129441. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26053163.

Ribas, L. G. dos S., Pressey, R. L., Loyola, R. & Bini, L. M. (2020) A global comparative analysis of impact evaluation methods in estimating the effectiveness of protected areas. *Biological Conservation*. 246, 108595. Available from: http://www.sciencedirect.com/science/article/pii/S0006320719319032.

Robalino, J., Pfaff, A. & Villalobos, L. (2017) Heterogeneous Local Spillovers from Protected Areas in Costa Rica. *Journal of the Association of Environmental and Resource Economists*. 4(3), 795–820. Available from: https://doi.org/10.1086/692089.

Rodriguez, E., Morris, C. & Belz, E. (2006) A global assessment of SRTM performance. *Photogrammetric Engineering and Remote Sensing*. 72, 249–260.

Rodríguez, N., Armenteras, D. & Retana, J. (2013) Effectiveness of protected areas in the Colombian Andes: deforestation, fire and land-use changes. *Regional Environmental Change*. 13(2), 423–435. Available from: https://doi.org/10.1007/s10113-012-0356-8.

Rosenbaum, P. R. & Rubin, D. B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*. 70(1), 41–55. Available from: https://doi.org/10.1093/biomet/70.1.41.

RStudio Team (2016). *RStudio: Integrated Development for R*. RStudio: PBC, Boston, MA Available from: http://www.rstudio.com/

Rubin, D. B. (1980) Bias Reduction Using Mahalanobis-Metric Matching. *Biometrics*. 36(2), 293–298. Available from: http://www.jstor.org/stable/2529981.

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

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V & Woolmer, G. (2002) The Human Footprint and the Last of the Wild: The human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. *BioScience*. 52(10), 891–904. Available from: https://doi.org/10.1641/0006-3568(2002)052[0891:THFATL]2.0.CO.

Sannigrahi, S., Bhatt, S., Rahmat, S., Paul, S. K. & Sen, S. (2018) Estimating global ecosystem service values and its response to land surface dynamics during 1995–2015. *Journal of Environmental Management*. 223, 115–131. Available from: http://www.sciencedirect.com/science/article/pii/S0301479718306297.

Sayer, J. (1991) *Rainforest Buffer Zones: Guidelines for Protected Area Managers*. IUCN World Conservation Union Forest Conservation Programme: Gland, Switzerland. p2. Available from: https://portals.iucn.org/library/sites/library/files/documents/FR-001.pdf

Schleicher, J., Eklund, J., D. Barnes, M., Geldmann, J., Oldekop, J. A. & Jones, J. P. G. (2020) Statistical matching for conservation science. *Conservation Biology*. 34(3), 538–549. Available from: https://doi.org/10.1111/cobi.13448.

Schwarze, R., Niles, J. O. & Olander, J. (2002) Understanding and managing leakage in forest–based greenhouse–gas–mitigation projects. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*. 360(1797), 1685–1703. Available from: https://doi.org/10.1098/rsta.2002.1040.

Sekhon, J. (2011) Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*. 42(7).

Serrano, F., Cardoza, G., Posadas, O., Hernández, E., González Cerdeira, Y., Landa, P., Mejía Rivera, J., Ayala, M. (2016) *Impacto socioambiental de la minería en la región noroccidental de Honduras a la luz de tres estudios de casos: Montaña de Botaderos (Aguán),Nueva Esperanza (Atlán da) y Locomapa (Yoro)*. Equipo de Reflexión, Investigación y Comunicación (ERIC) & Saint Louis University: Yoro, Honduras. Available from: http://www.acafremin.org/images/documentos/E-Book_Informe-industria-extractivaVF-8-7-16.pdf

Shah, P. & Baylis, K. (2015) Evaluating Heterogeneous Conservation Effects of Forest Protection in Indonesia. *PLOS ONE*. 10(6), e0124872. Available from:

https://doi.org/10.1371/journal.pone.0124872.

Spracklen, B. D., Kalamandeen, M., Galbraith, D., Gloor, E. & Spracklen, D. V (2015) A Global Analysis of Deforestation in Moist Tropical Forest Protected Areas. *PLOS ONE*. 10(12), e0143886. Available from: https://doi.org/10.1371/journal.pone.0143886.

Stuart, E. A. (2010) Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*. 25(1), 1–21. Available from: https://pubmed.ncbi.nlm.nih.gov/20871802.

UNEP-WCMC (2019). *User Manual for the World Database on Protected Areas and world database on other effective area-based conservation measures:1.6*. UNEP-WCMC: Cambridge, UK. Available from: http://wcmc.io/WDPA_Manual

UNEP-WCMC (2020). *August 2020 update of the WDPA* [Online], Cambridge, UK: UNEP-WCMC. Available from: https://www.protectedplanet.net/c/monthly-updates/2020/august-2020-update-of-the-wdpa [Accessed August 2020]

UNEP-WCMC and IUCN (2020a), *Protected Planet: The World Database on Protected Areas (WDPA)* [Online], Cambridge, UK: UNEP-WCMC and IUCN. Available from: www.protectedplanet.net. [Accessed June 2020]

U.S. Geological Survey (USGS) Earth Resources Observation and Science Center (EROS) (2014*) SRTM3 Collection*. U. S. Geological Survey (USGS) Earth Resources Observation and Science Center (EROS): Sioux Falls, South Dakota. [Accessed July 2020]

Venter, O., Magrach, A., Outram, N., Klein, C. J., Possingham, H. P., Di Marco, M. & Watson, J. E. M. (2018) Bias in protected-area location and its effects on long-term aspirations of biodiversity conventions. *Conservation Biology*. 32(1), 127–134. Available from: https://doi.org/10.1111/cobi.12970.

Wang, C. & Myint, S. W. (2016) Environmental Concerns of Deforestation in Myanmar 2001–2010. *Remote Sensing*. 8 (9).

Watson, J. E. M., Dudley, N., Segan, D. B. & Hockings, M. (2014) The performance and potential of protected areas. *Nature*. 515(7525), 67–73. Available from: https://doi.org/10.1038/nature13947.

Wittemyer, G., Elsen, P., Bean, W. T., Burton, A. C. O. & Brashares, J. S. (2008) Accelerated Human Population Growth at Protected Area Edges. *Science*. 321(5885), 123–126. Available from: http://science.sciencemag.org/content/321/5885/123.abstract.

Yang, H., Viña, A., Winkler, J. A., Chung, M. G., Dou, Y., Wang, F., Zhang, J., Tang, Y., Connor, T., Zhao, Z. & Liu, J. (2019) Effectiveness of China's protected areas in reducing deforestation. *Environmental Science and Pollution Research*. 26(18), 18651–18661. Available from: https://doi.org/10.1007/s11356-019-05232-9.

Zalasiewicz, J., Williams, M., Steffen, W. & Crutzen, P. (2010) The New World of the Anthropocene. *Environmental Science & Technology*. 44(7), 2228–2231. Available from: https://doi.org/10.1021/es903118j.

Zhang, H. & Malczewski, J. (2017) Accuracy Evaluation of the Canadian OpenStreetMap Road Networks. *International Journal of Geospatial and Environmental Research*. 5(2). Available from: https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1364&context=geographypub.

Appendices

Appendix A

Table 1. Selection of papers assessing deforestation in PAs demonstrating the typical study time to generate deforestation rates.

Paper	Duration
Curran et al (2004)	1988-2002
Sanchez-Azofeifa et al. (2003)	1960–1979, 1979–1986 and 1986–1997
Oliveira et al. (2007)	1999-2005
Andam et al. (2008)	1955-1960 and 1986-1997
Gaveau et al. (2009)	1990-2000
Joppa and Pfaff (2011)	2000-2005
Nelson et al. (2011)	1990-2000
Rodriguez et al. (2013)	1985-2005
Heino et al. (2015)	2000-2012
Rasolofoson et al. (2015)	≥5 years
Spracklen et al (2015)	2000-2012
Lui and Coombes (2016)	2000-2012
Alix-Garcia and Gibbs (2017)	2007-2015
Herrera et al. (2019)	2000-2004 & 2004-2008
Oldekop et al. (2019)	2000-2012
Poor et al. (2019)	2002-2016
Yang et al. (2019)	2000-2012

Appendix B

Table 1. All data layers, source, type, and date.

Data	Source	Type/ Resolution	Year
Colombian administrative areas, roads,	National Administrative Department of Statistics (National Administrative Department of Statistics, 2018): Accessed June 2020 from https://geoportal.dane.gov.co/	Vector	2018
Peruvian administrative areas, roads	Derived from Humanitarian OpenStreetMap Team* (HOT): Accessed July 2020 from https://data.humdata.org	Vector	2020
Honduran administrative area, roads	Derived from Humanitarian OpenStreetMap Team* (HOT): Accessed July 2020 from https://data.humdata.org	Vector	2019
Cameroonian administrative areas	Institut National de Cartographie (INC): Accessed July 2020 from https://data.humdata.org/dataset/cameroon-administrative-boundaries	Vector	2019
Cameroonian roads	Derived from Humanitarian OpenStreetMap Team* (HOT): Accessed July 2020 from https://data.humdata.org/	Vector	2020
Indian administrative areas, roads	Derived from Humanitarian OpenStreetMap Team* (HOT): Accessed July 2020 from https://data.humdata.org/	Vector	2020
Philippines administrative area, roads	Derived from Humanitarian OpenStreetMap Team* (HOT): Accessed July 2020 from https://data.humdata.org	Vector	2020
Protected area boundaries and metadata	World Database on Protected Areas (UNEP-WCMC and IUCN, 2020): Accessed June 2020 from https://www.protectedplanet.net/	Vector	2020
Forest cover (2000) & Deforestation (2001-2019)	Global Forest Change v1. 7 (Hansen *et al*., 2013): Accessed June 2020 from https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.7.html	Raster / 30m	2000 - 2019
Elevation	Shuttle Radar Topography Mission (SRTM) (USGS, 2014): Accessed June 2020 using QGIS plugin SRTM Downloader v3.1.4	Raster / 30m	2000
Slope	Generated using QGIS slope command from the SRTM digital elevation model	Raster / 30m	2000
Ecoregion	WWF Terrestrial ecoregions of the world (Olson et al., 2001): Accessed June 2020 from https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world	Vector	2001
Urban settlements	GRUMP (Center for International Earth Science Information Network *et al*., 2017) Accessed June 2020 from https://sedac.ciesin.columbia.edu/data/set/grump-v1-settlement-points-rev01/data-download	Vector	2000

*All OpenStreetMap data copyright OpenStreetMap contributors and available from https://www.openstreetmap.org

Appendix C

Table 1. Projected Coordinate Reference System used for each country in the GIS analysis.

Name	WDPA ID	Country	Coordinate reference system (CRS)	Justification
Boumba Bek/Nki	308624 & 30674	Cameroon	WGS84 / UTM zone 33N (EPSG: 32633)	The current Cameroon geodetic network is based on the WGS84 (Kande et al., 2016)
Deng Deng	555547995			
San Miguel de los Farallones	555555800	Colombia	MAGNA-SIRGAS / Colombia East Central Zone (EPSG:3117)	Standard projection used by the Colombian administration (Martinez and Sanchez, 2009)
Montaña de Botaderos Carlos Escaleras Mejía	555582981	Honduras	WGS84 / UTM zone 16N (EPSG: 32616)	Honduras does not have a dedicated CRS so the appropriate WGS UTM zone was used.
Congolón, Piedra Parada y Coyocutena	62051			
Kyauk Pan Taung	1235	Myanmar	WGS84 / UTM zone 47 (Myanmar Datum 2000)	The Myanmar datum was developed and adopted by the Myanmar Ministry of Natural Resources and Environmental Conservation, Survey Department (Aung Moe, 2016)
Papikonda	1774	India	WGS84 / India NSF LCC (EPSG:7755)	WGS84 has been adopted by the Indian government since the 2005 National Map Policy (Ghosh and Dubey, 2008)
Bosques Nublados de Udima	555544103	Peru	PSAD56 / Peru Central Zona (EPSG: 24982)	Developed by the Peruvian administration covering the study area (Mugnier, 2006)
Mount Balatukan Range	555583087	Philippines	Philippine Reference System 1992 (PRS 92) Philippines zone 5 (EPSG: 3125)	This is the national geodetic system developed by the Phillipines National Mapping and Resource Information Authority (NAMRIA) (Fajardo, 2001)

Appendix D

Table 1. Standardised mean difference between pre- and post-matched buffer and PA samples for all 9 PAs analysed. Values highlighted in red are above the acceptable level of difference (0.25), green values are below the acceptable level, and values in bold green are below the desired level of difference (0.1).

Covariate	Boumba Bek/Nki (Cam)				Deng Deng (Cam)				San Miguel de los Farallones (Col)				Congolón, Piedra Parada y Coyocutena (Hon)				Montaña de Botaderos Carlos Escaleras Mejía (Hon)			
	PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff	
	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
Distance to settlement	2.96	**0.02**	2.03	**0.02**	1.51	**0.05**	0.91	**0.03**	4.16	0.19	3.33	**0.08**	4.94	**0.04**	1.75	**0.05**	0.35	**0.04**	0.11	**0.00**
Distance to road	1.76	0.11	0.60	**0.04**	0.30	**0.01**	0.40	**0.07**	0.59	**0.09**	1.72	**0.05**	1.75	**0.07**	0.15	**0.02**	0.46	**0.03**	0.14	**0.03**
Slope	0.22	0.11	0.24	**0.04**	0.27	**0.03**	0.31	**0.02**	0.27	0.18	0.11	**0.03**	0.38	0.10	0.25	**0.01**	0.72	**0.02**	0.24	**0.02**
Elevation	2.05	**0.02**	2.30	0.13	0.14	**0.01**	0.99	**0.01**	2.37	**0.05**	0.69	**0.03**			1.23	0.10	0.91	**0.06**	0.10	**0.03**
Distance to forest edge	1.08	**0.06**	0.58	**0.06**	0.29	**0.01**	0.20	**0.01**	2.06	0.17	1.69	**0.08**	0.82	**0.01**	0.67	**0.02**	0.79	**0.03**	0.36	**0.00**
Canopy cover	6.30	0.15	6.15	**0.03**	1.50	**0.05**	1.17	**0.02**	**0.01**	0.12	0.25	**0.07**	0.30	0.11	0.16	**0.06**	2.19	**0.09**	0.73	**0.00**

Covariate	Papikonda (Ind)				Kyauk Pan Taung (Mya)				Bosques Nublados de Udima (Per)				Mount Balatukan Range (Phi)			
	PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff		PA Std. Mean diff		Buffer Std. Mean diff	
	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
Distance to settlement	**0.05**	**0.04**	0.11	**0.00**	10.20	**0.02**	3.47	**0.02**	0.57	**0.00**	0.37	**0.03**				
Distance to road	0.36	**0.02**	0.29	**0.02**	0.14	**0.06**	0.48	**0.03**	**0.05**	**0.03**	0.31	**0.01**	0.71	**0.05**	1.25	**0.09**
Slope	0.64	**0.02**	0.40	**0.01**	0.56	0.18	0.49	0.10	**0.05**	**0.01**	0.34	**0.01**	0.90	**0.06**	0.34	**0.02**
Elevation	0.74	**0.07**	1.11	**0.00**	0.80	0.11	2.56	**0.04**	0.64	**0.02**	0.53	**0.04**	2.71	**0.08**	0.78	**0.06**
Distance to forest edge	0.91	**0.05**	0.71	**0.04**	**0.10**	**0.07**	0.29	**0.05**	1.04	**0.03**	0.39	**0.05**	0.40	**0.04**	0.23	**0.04**
Canopy cover	1.81	**0.07**	1.35	**0.05**	0.36	**0.08**	0.31	**0.01**	1.38	**0.07**	0.18	**0.09**	0.62	**0.01**	**0.10**	**0.04**

Appendix E

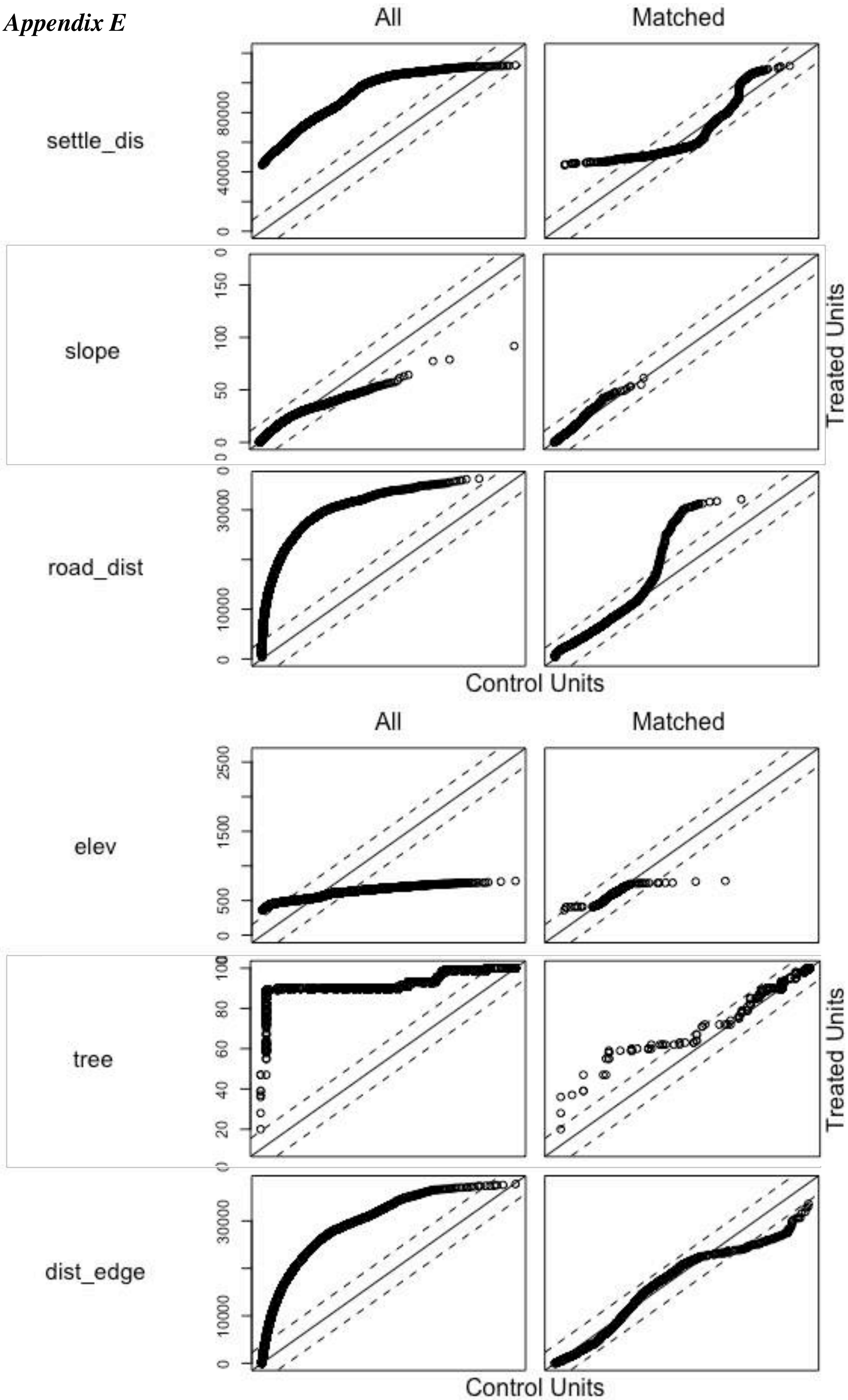

Figure 1. QQ plots of the covariate distributions for Boumba Bek/Nki (Cam) before and after matching. Matched samples N=1746.

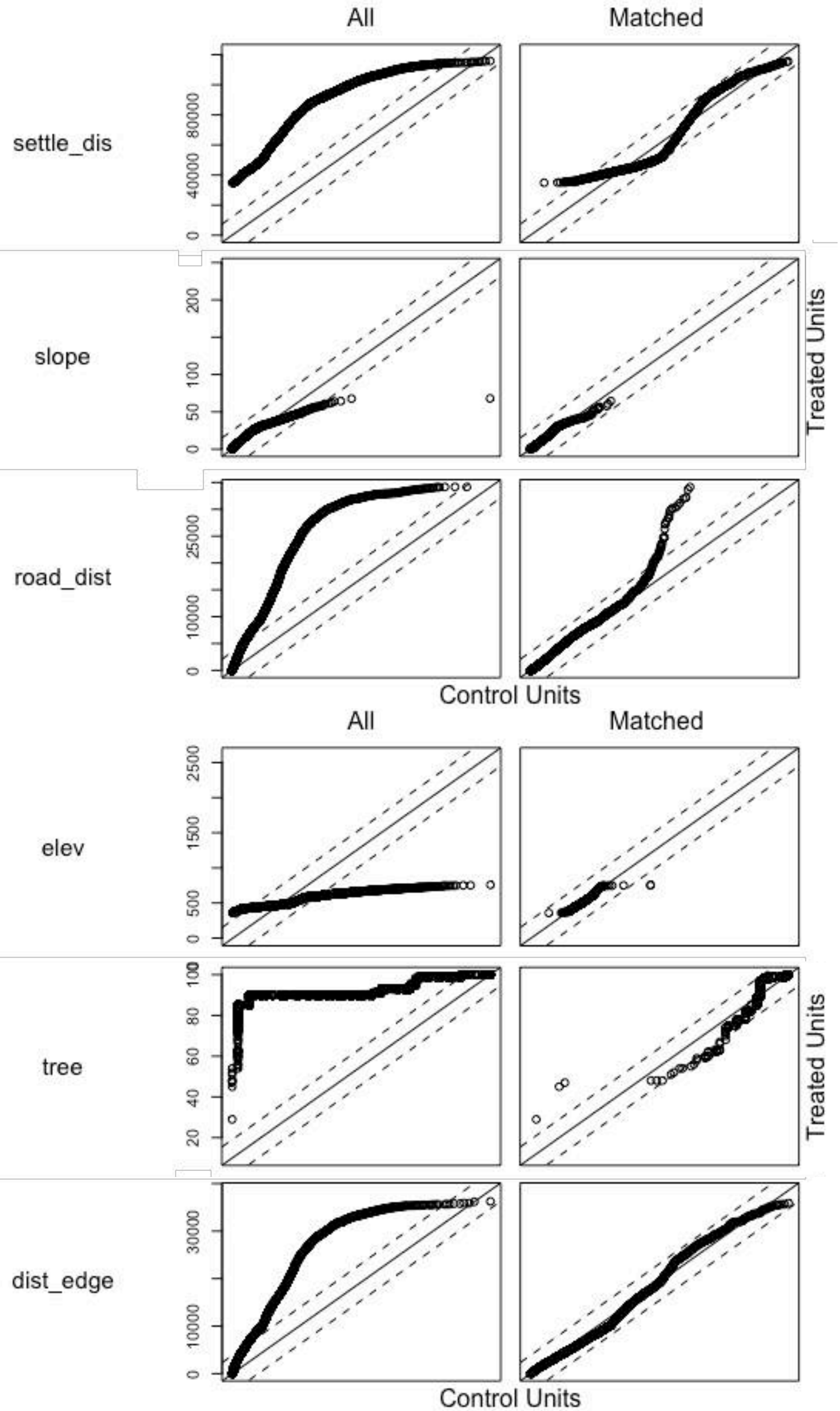

Figure 2. QQ plots of the covariate distributions for Boumba Bek/Nki (Cam) buffer before and after matching. Matched samples N=4301.

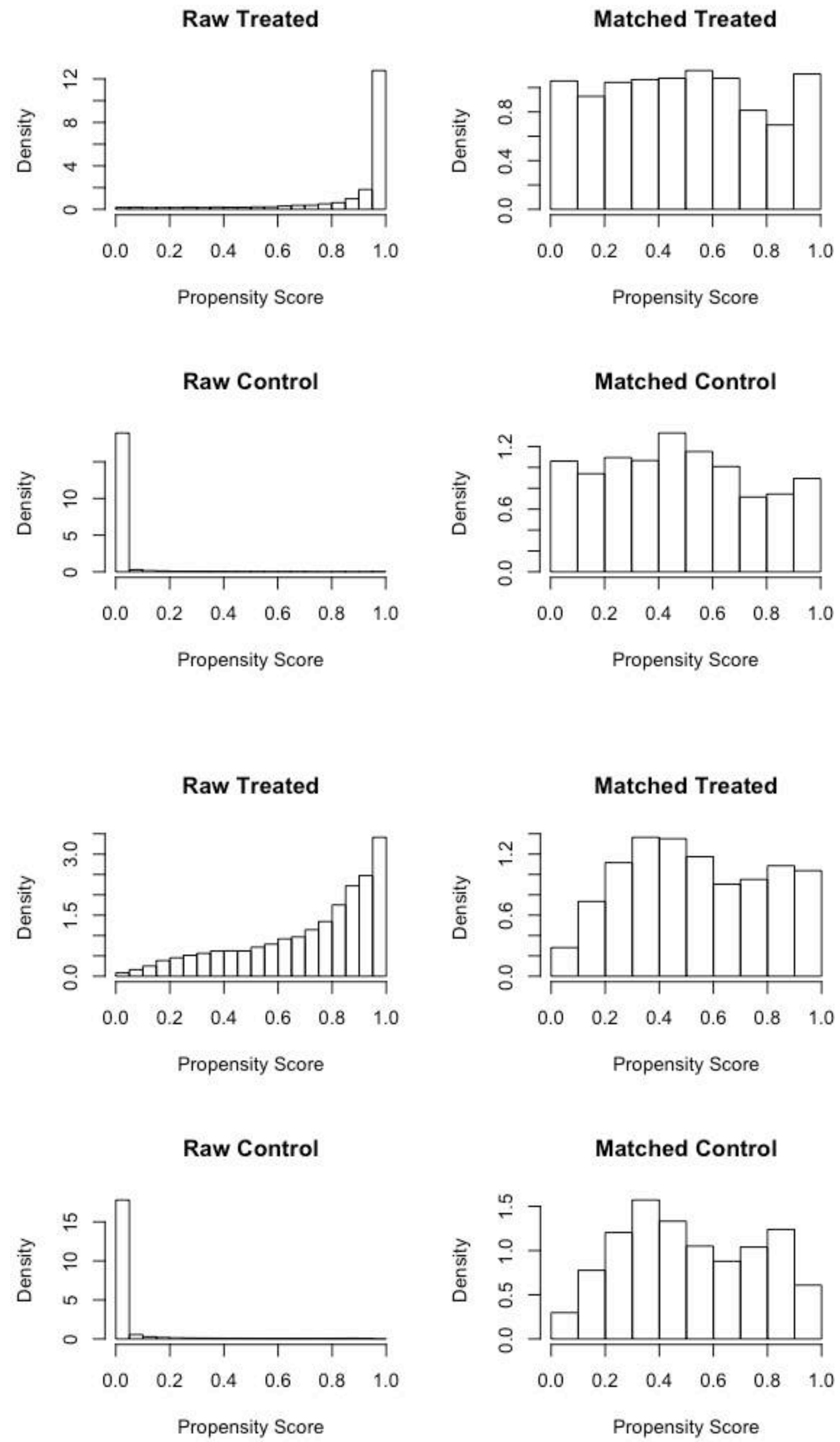

Figure 3. Histogram distributions of Boumba Bek/Nki (Cam) before and after matching for PA (top) and Buffer (bottom).

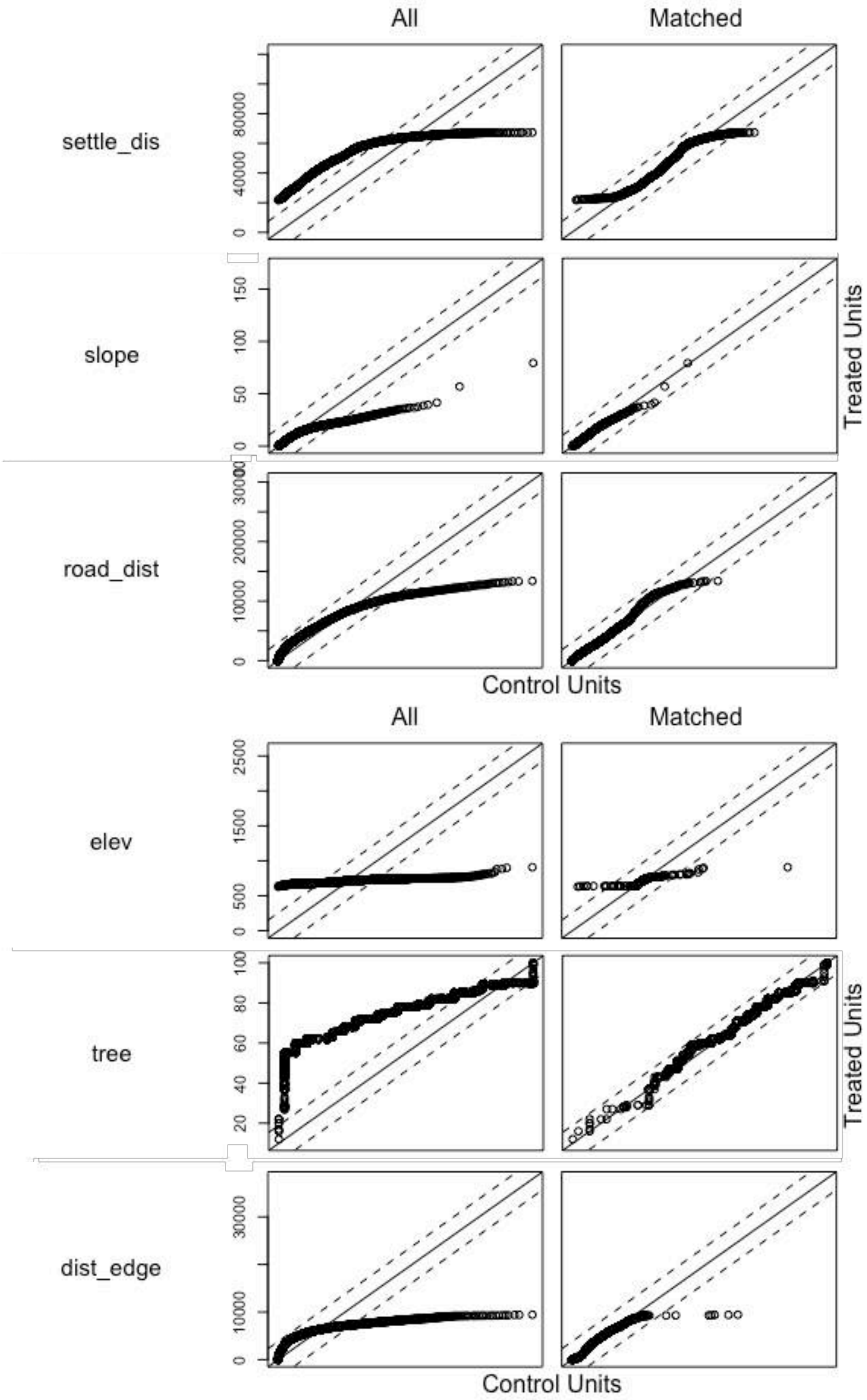

Figure 4. QQ plots of the covariate distributions for Deng Deng (Cam) before and after matching. Matched samples N=3874.

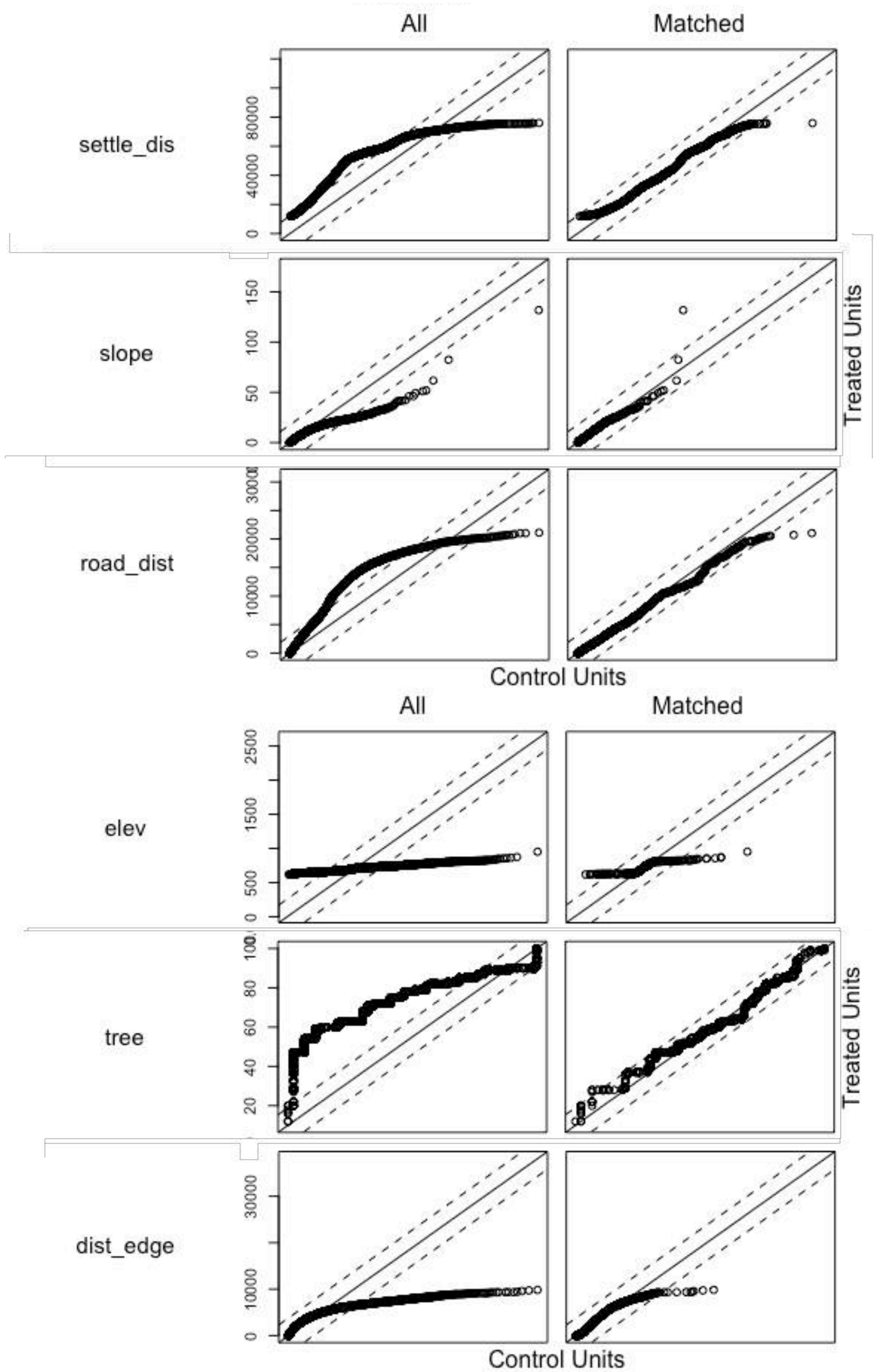

Figure 5. QQ plots of the covariate distributions for Deng Deng (Cam) buffer before and after matching. Matched samples N=5784.

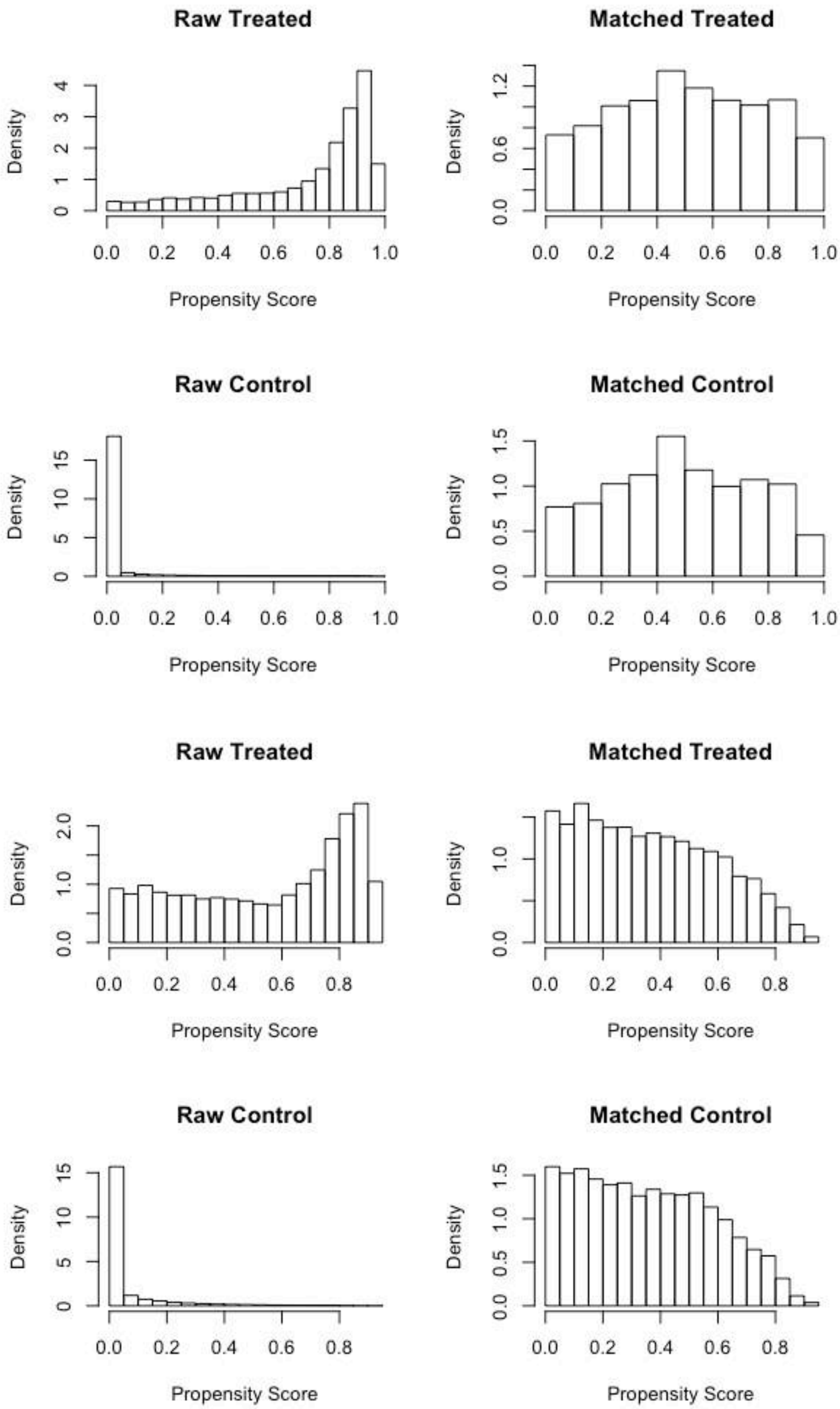

Figure 6. Histogram distributions of Deng Deng (Cam) before and after matching for PA (top) and Buffer (bottom).

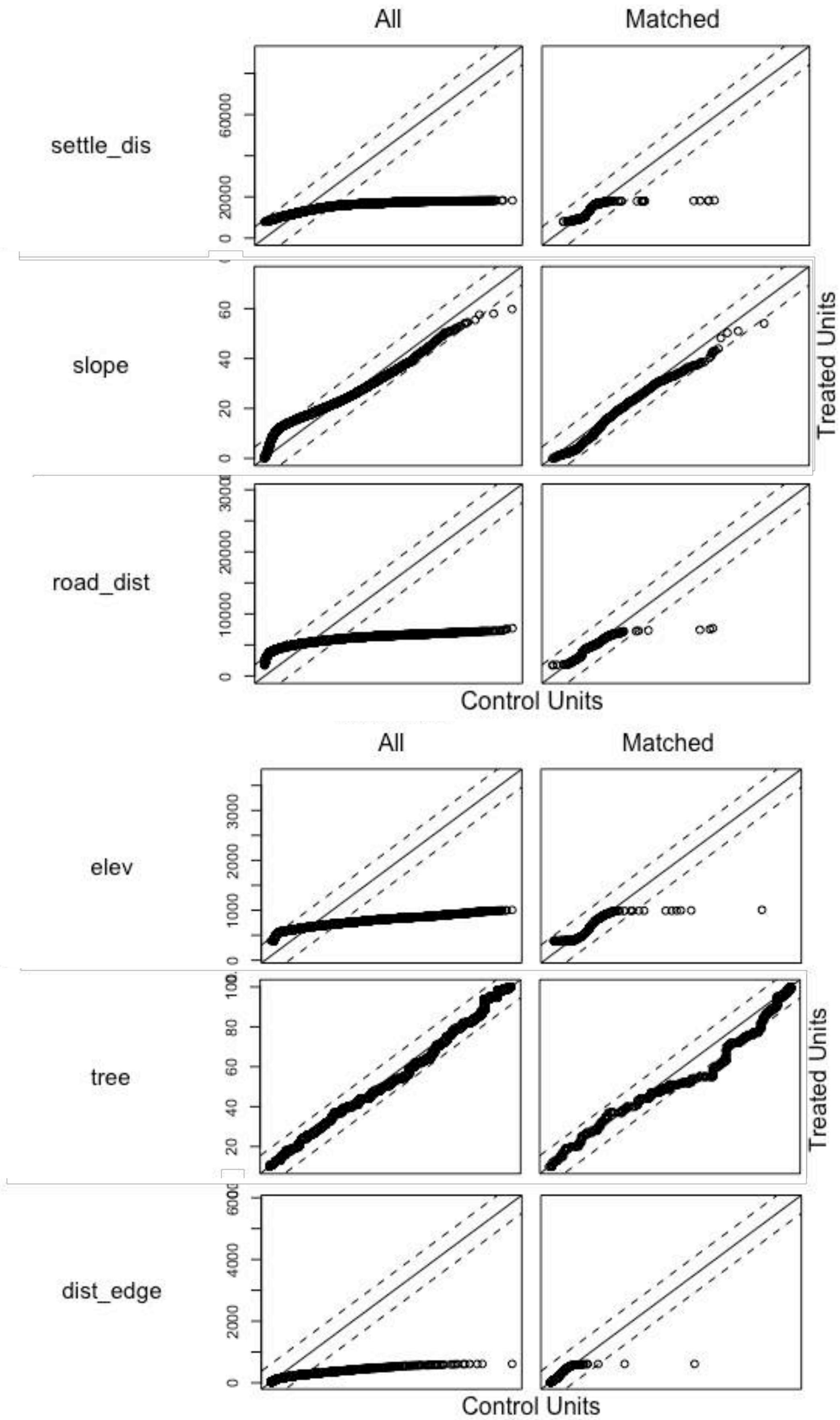

Figure 7. QQ plots of the covariate distributions for San Miguel de los Farallones (Col) before and after matching. Matched samples N=900

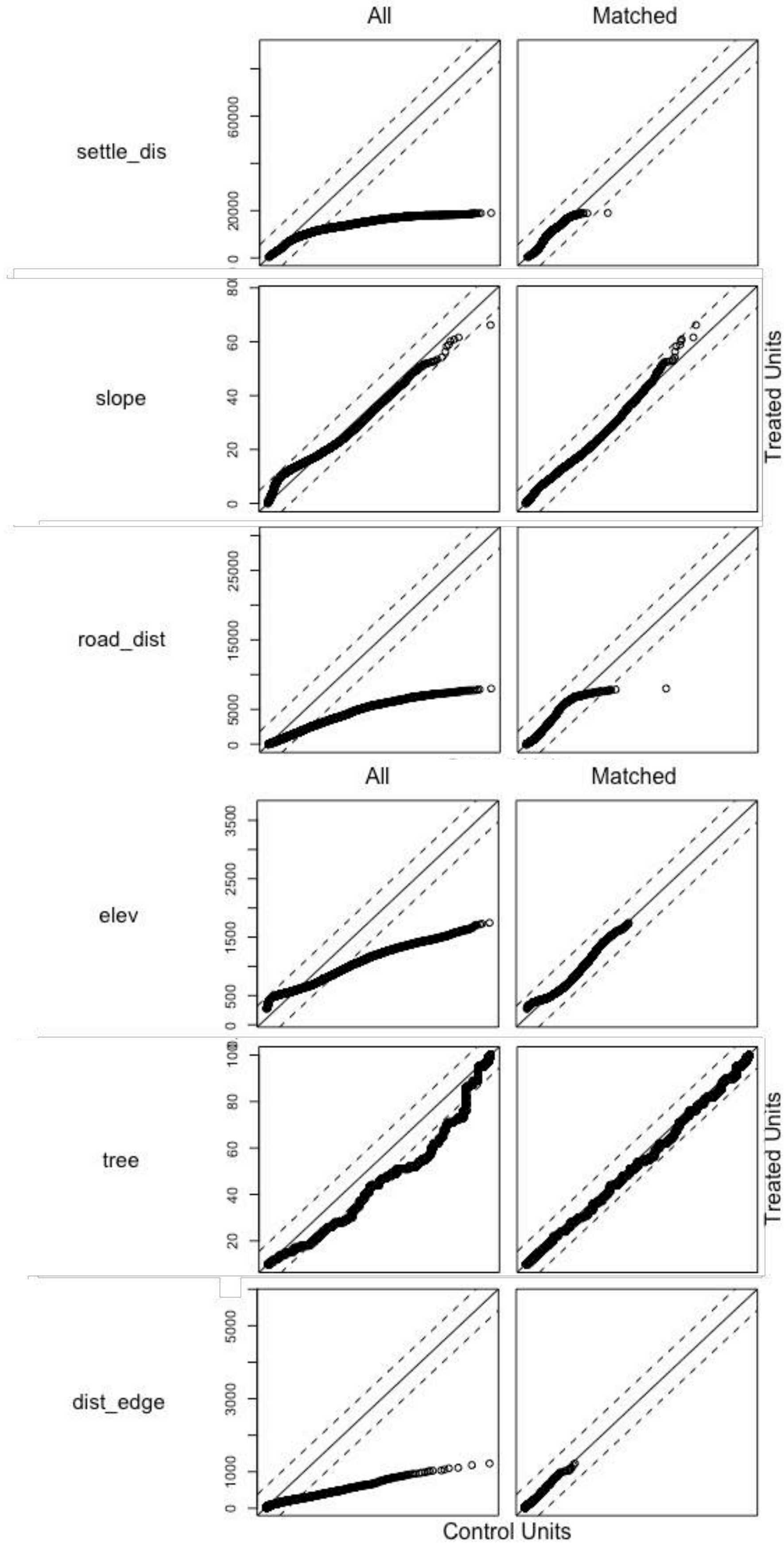

Figure 8. QQ plots of the covariate distributions for San Miguel de los Farallones (Col) buffer before and after matching. Matched samples N=3900

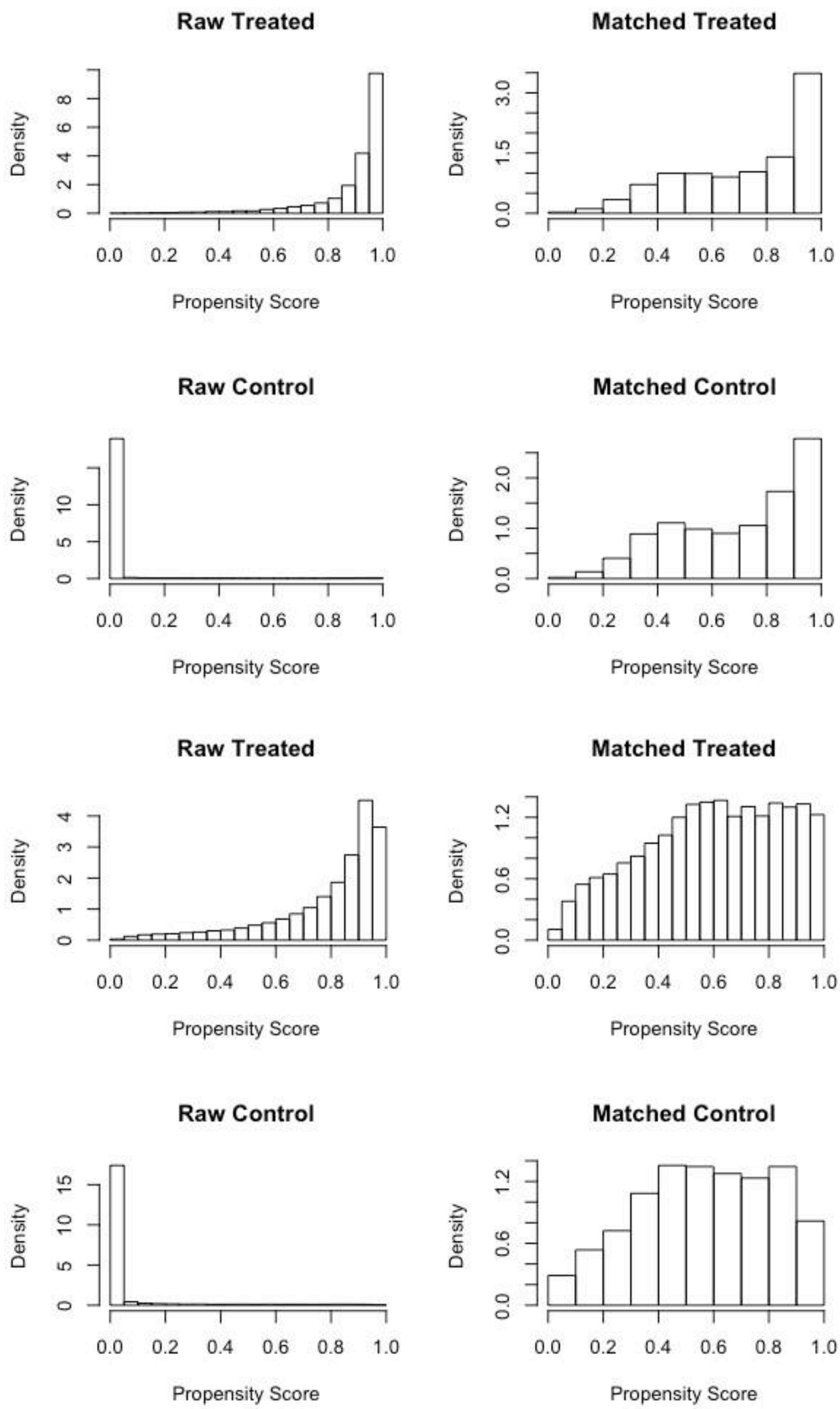

Figure 9. Histogram distributions of San Miguel de los Farallones (Col) before and after matching for PA (top) and Buffer (bottom).

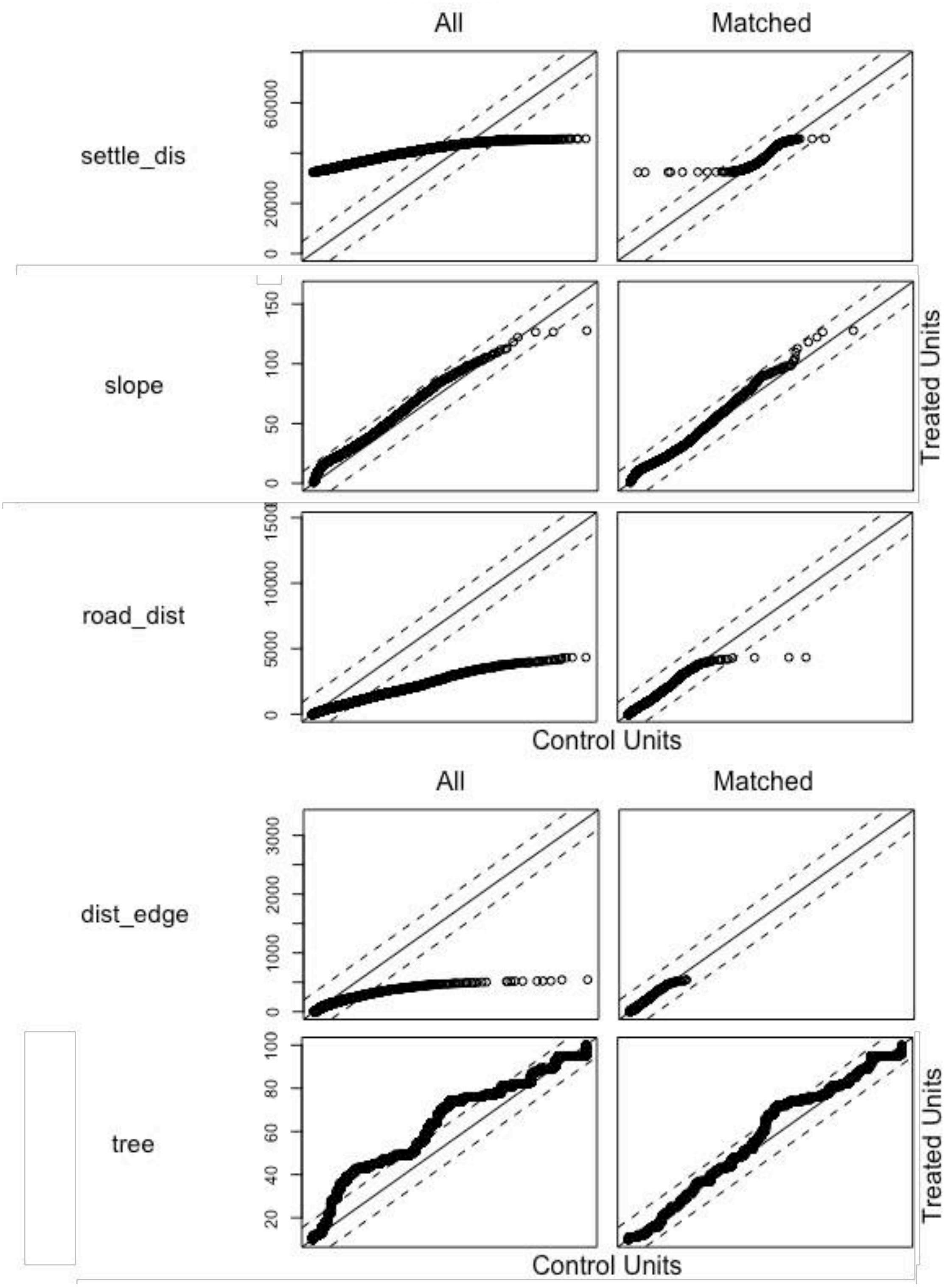

Figure 10. QQ plots of the covariate distributions for Congolón, Piedra Parada y Coyocutena (Hon) before and after matching. Note that elevation was excluded due to poor balance. Matched samples N=2827

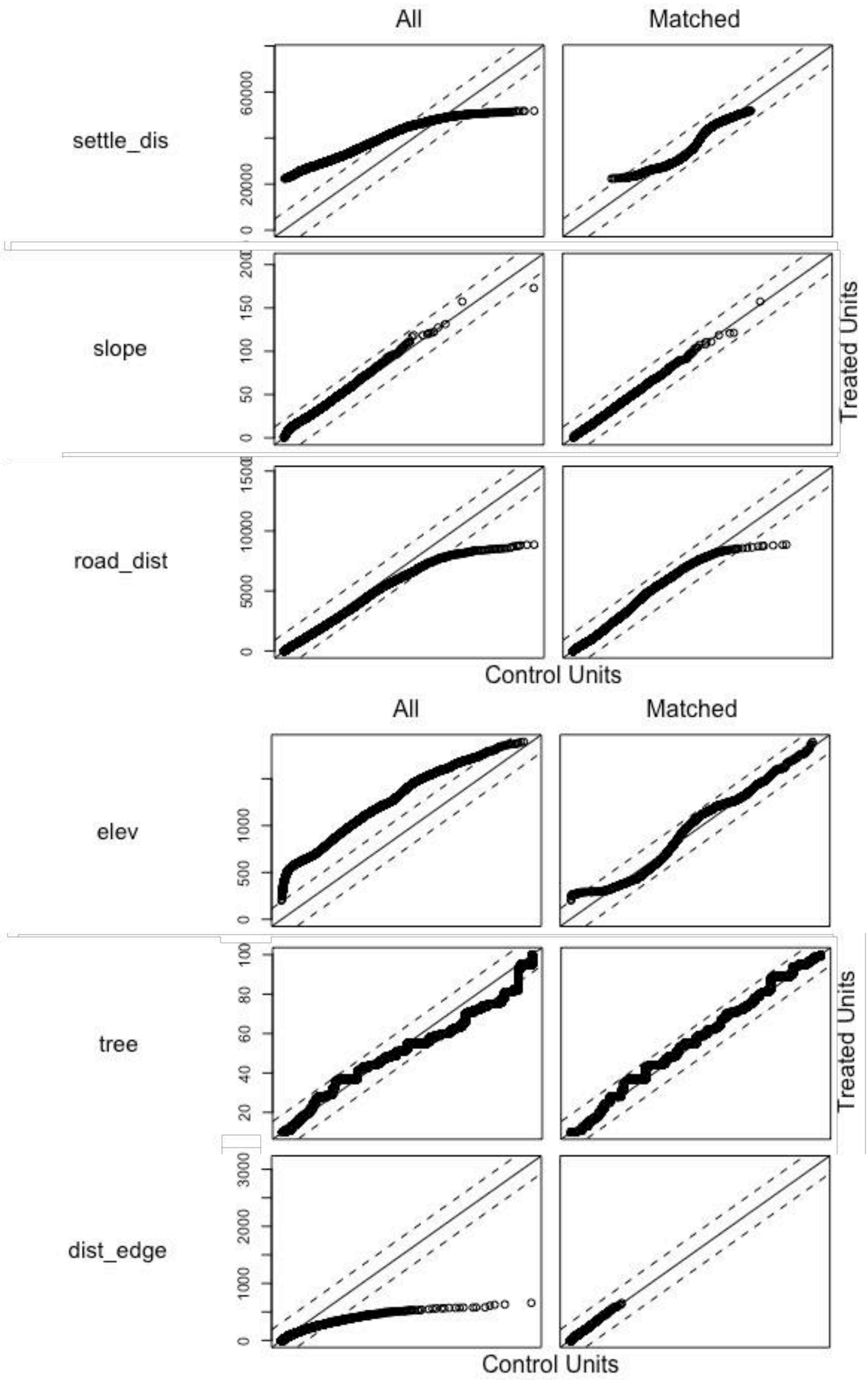

Figure 11. QQ plots of the covariate distributions for Congolón, Piedra Parada y Coyocutena (Hon) buffer before and after matching. Matched samples N=4706.

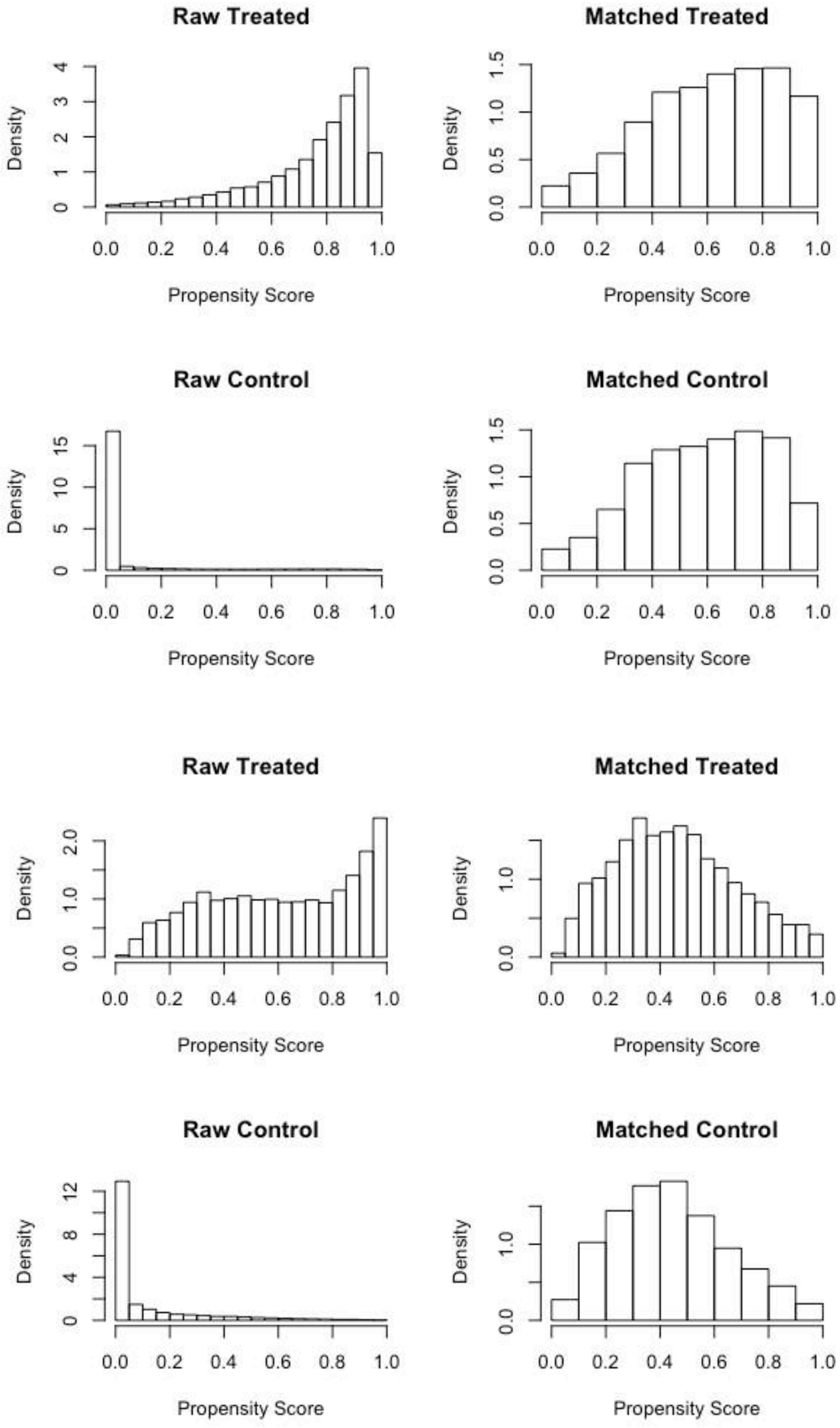

Figure 12. Histogram distributions of Congolón, Piedra Parada y Coyocutena (Hon) before and after matching for PA (top) and Buffer (bottom).

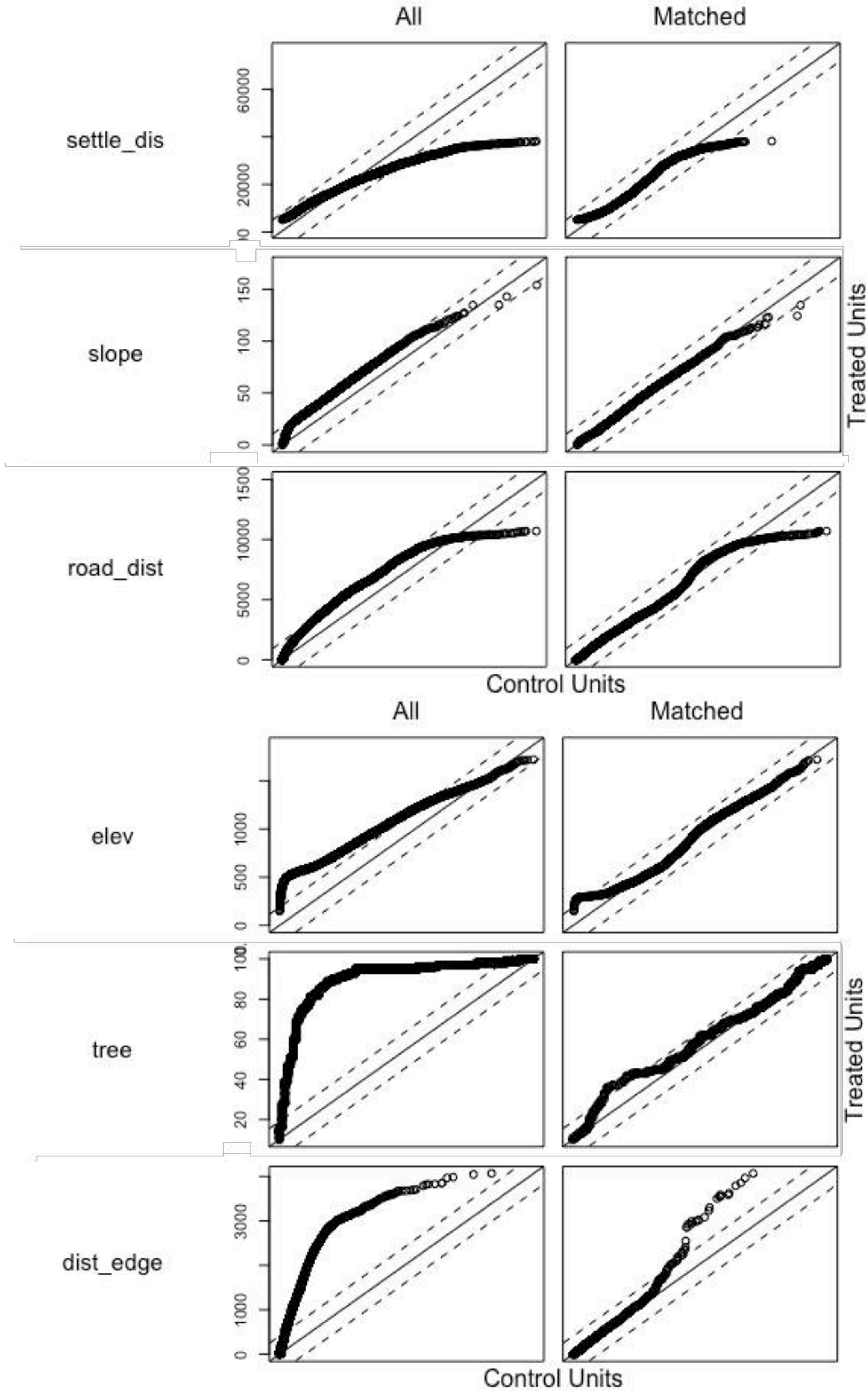

Figure 13. QQ plots of the covariate distributions for Montaña de Botaderos Carlos Escaleras (Hon) before and after matching. Matched samples N=5254

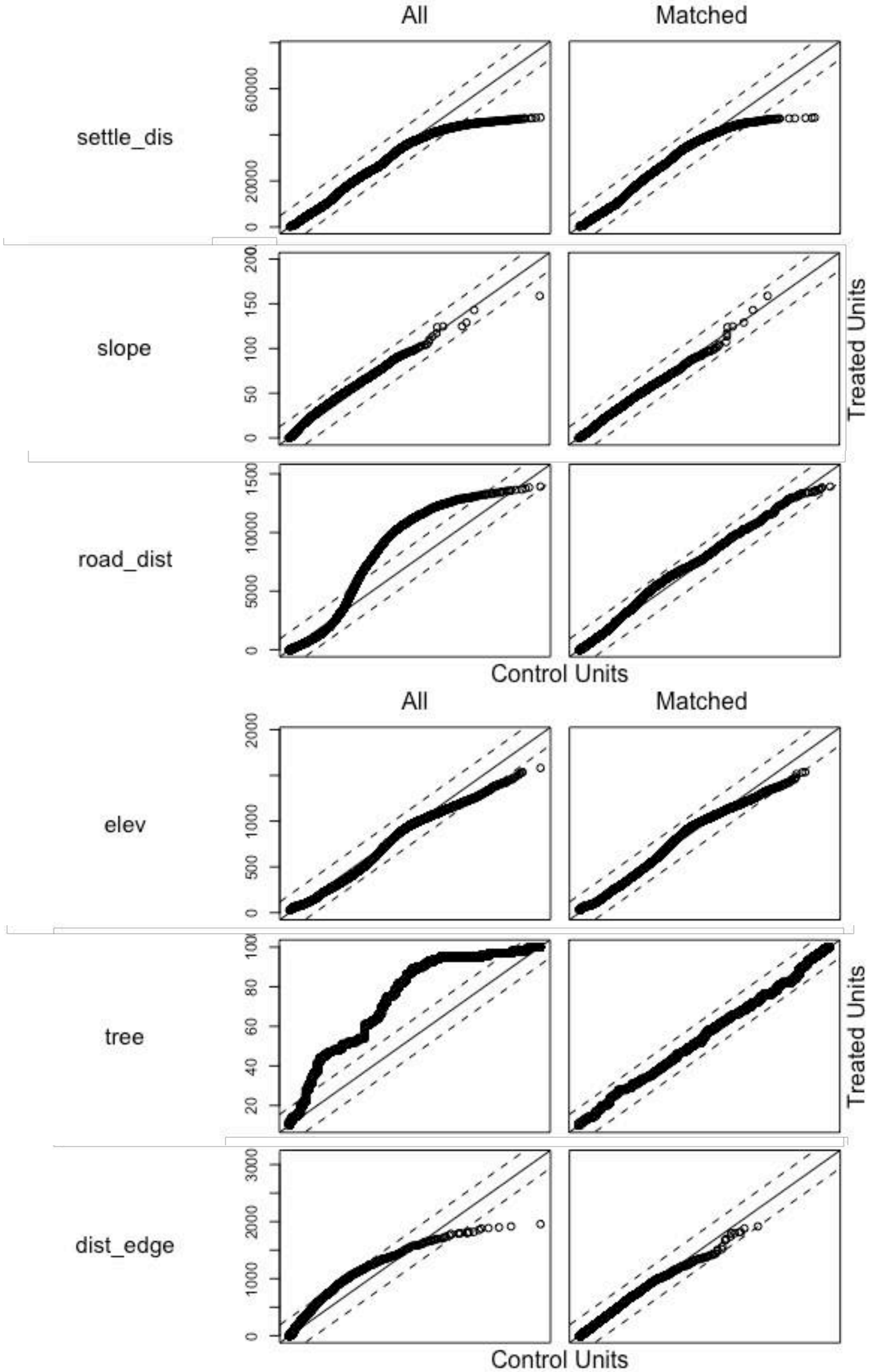

Figure 14. QQ plots of the covariate distributions for Montaña de Botaderos Carlos Escaleras (Hon) buffer before and after matching. Matched samples N=7988

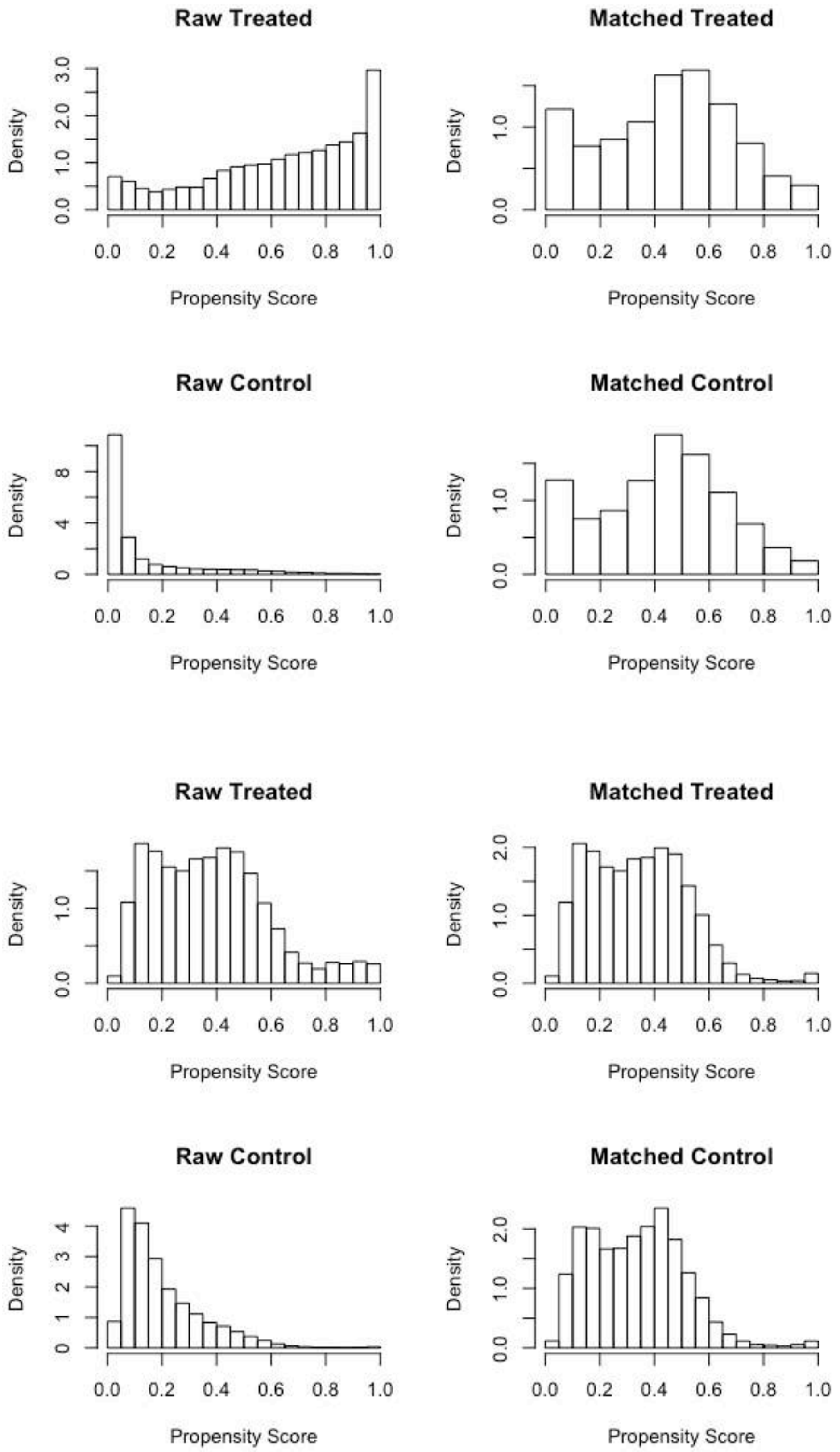

Figure 15. Histogram distributions of Montaña de Botaderos Carlos Escaleras (Hon) before and after matching for PA (top) and Buffer (bottom).

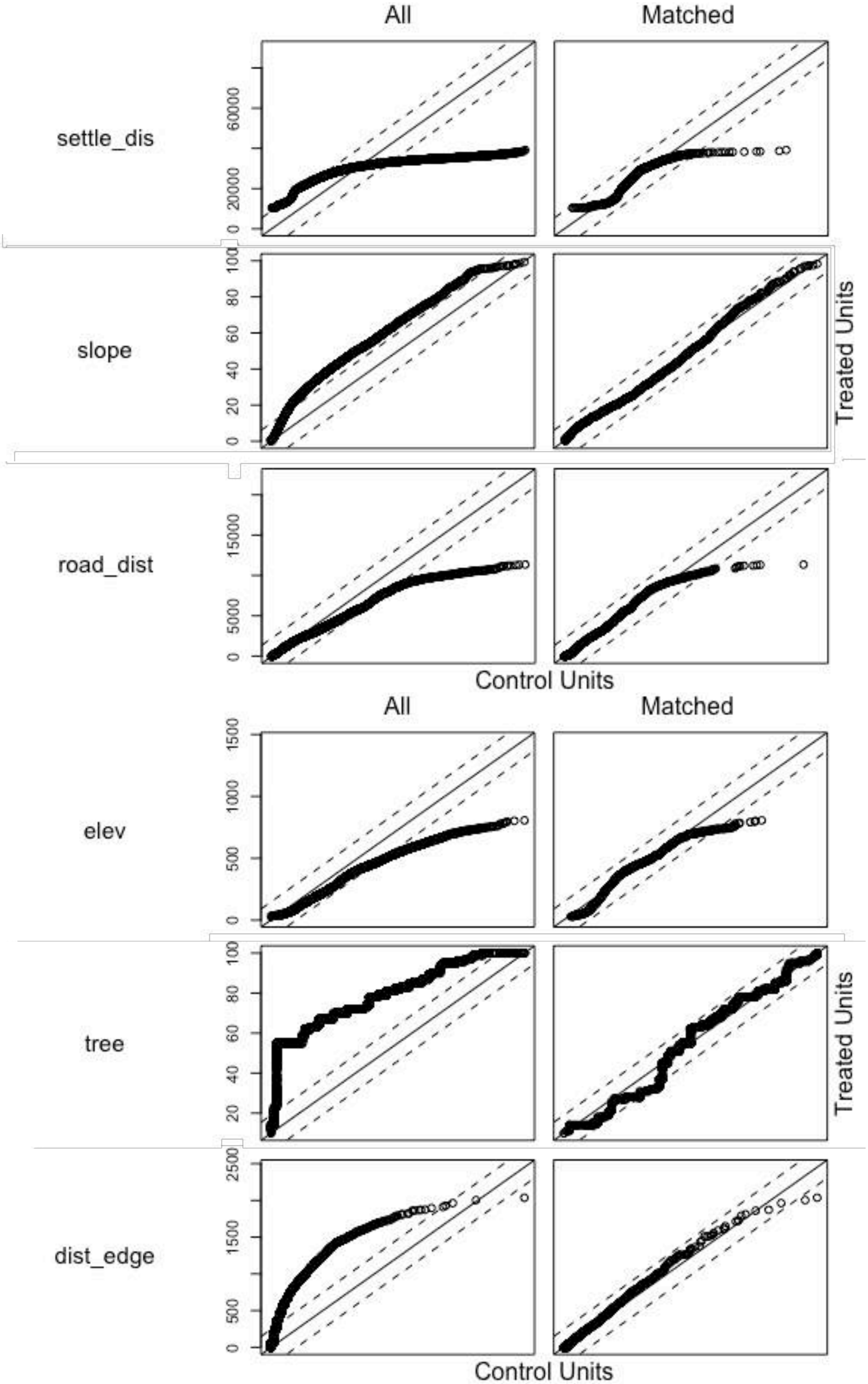

Figure 16. QQ plots of the covariate distributions for Papikonda (Ind) before and after matching. Matched samples N=2571

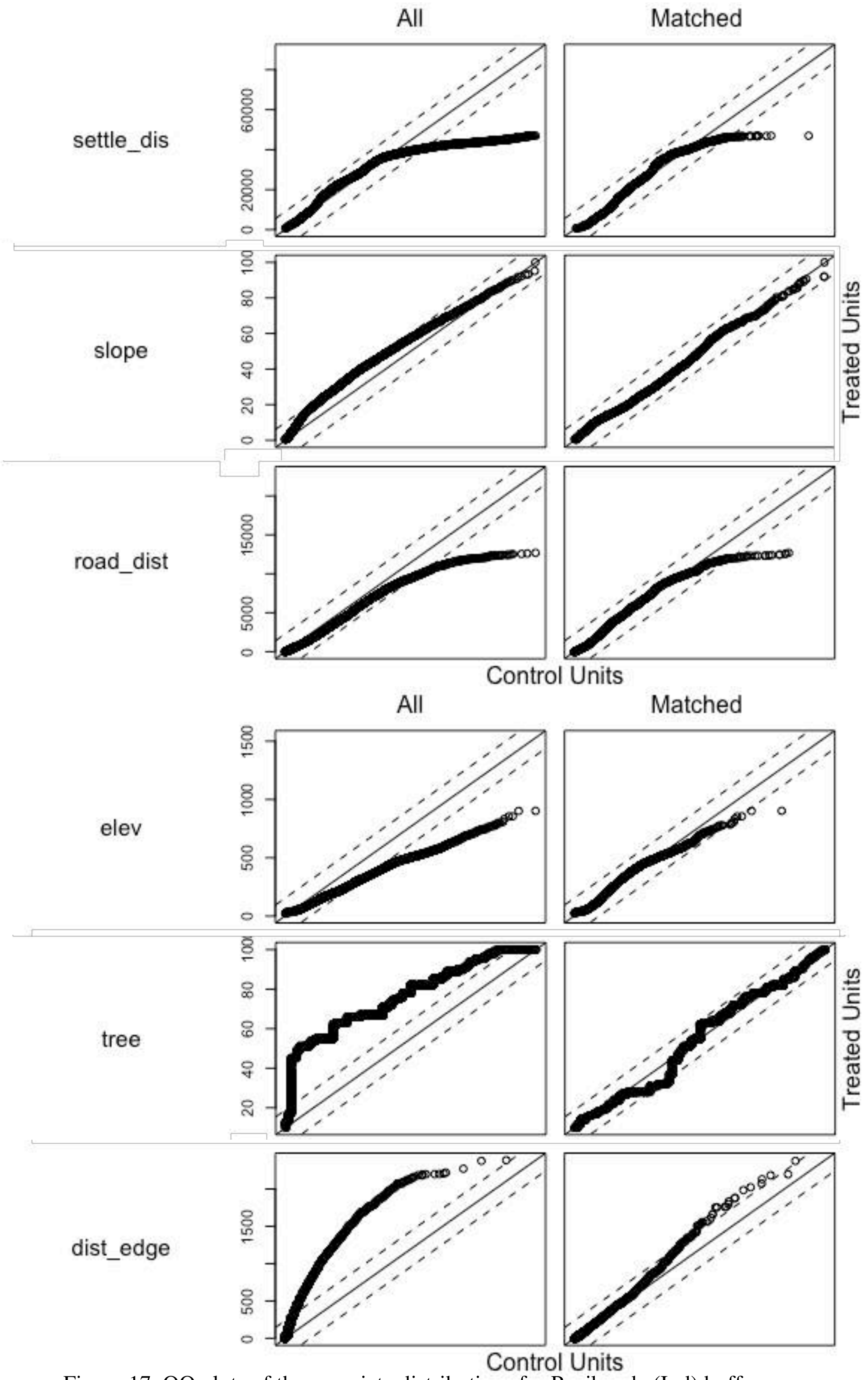

Figure 17. QQ plots of the covariate distributions for Papikonda (Ind) buffer before and after matching. Matched samples N=3628

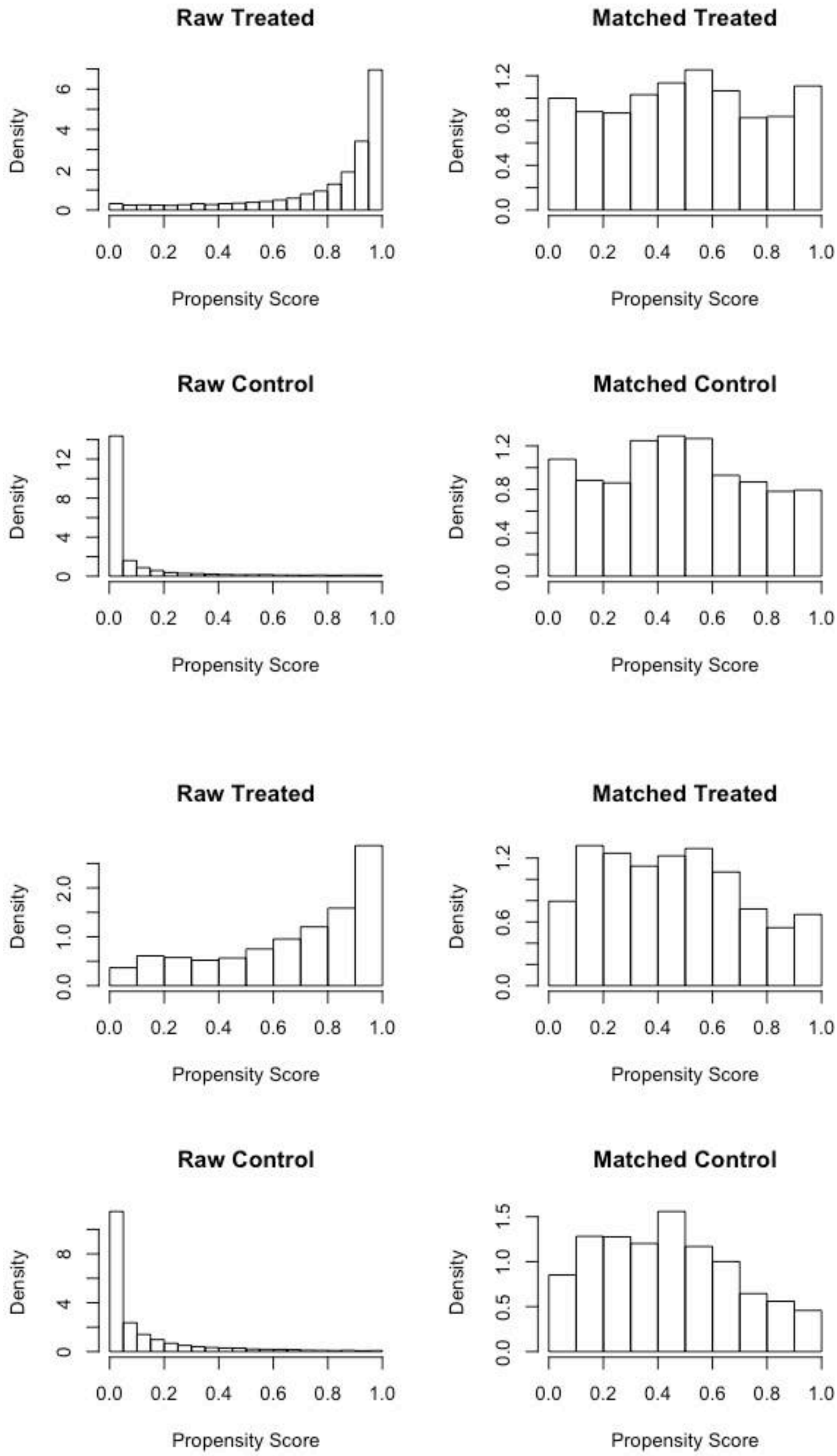

Figure 18. Histogram distributions of Papikonda (Ind) before and after matching for PA (top) and Buffer (bottom).

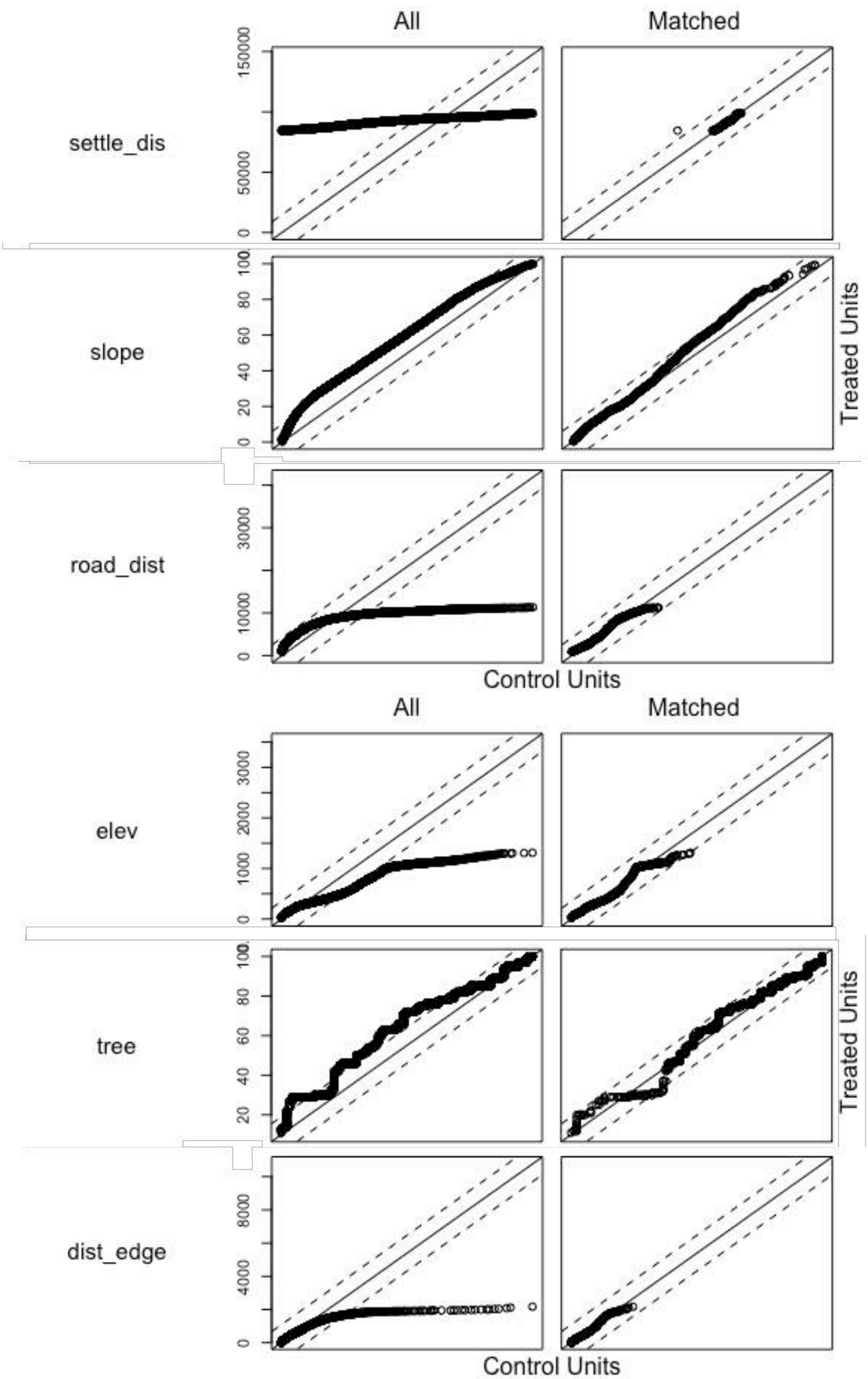

Figure 19. QQ plots of the covariate distributions for Kyauk Pan Taung (Mya) before and after matching. Matched samples N=1860

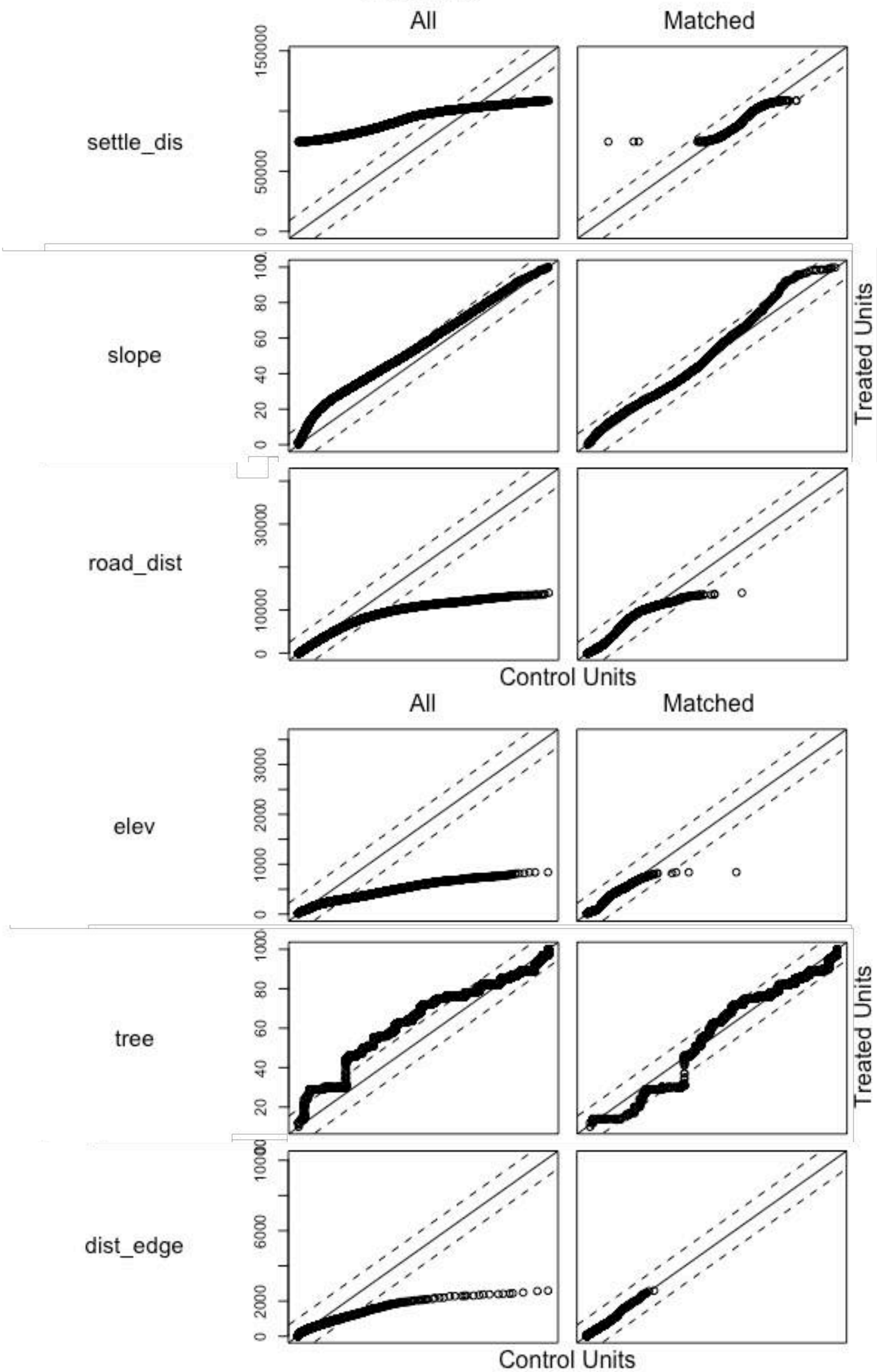

Figure 20. QQ plots of the covariate distributions for Kyauk Pan Taung (Mya) buffer before and after matching. Matched samples N=4583

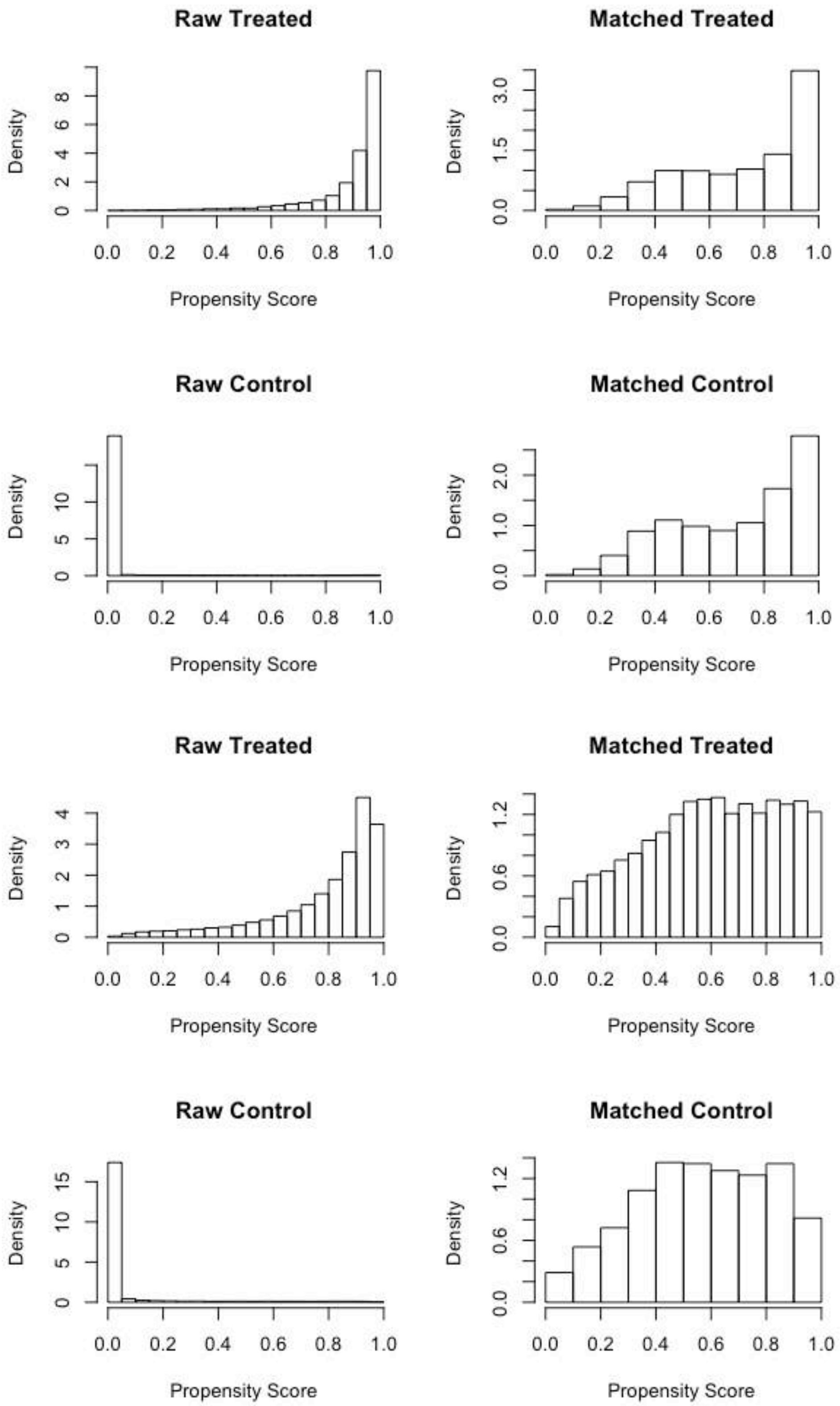

Figure 21. Histogram distributions of Kyauk Pan Taung (Mya) before and after matching for PA (top) and Buffer (bottom).

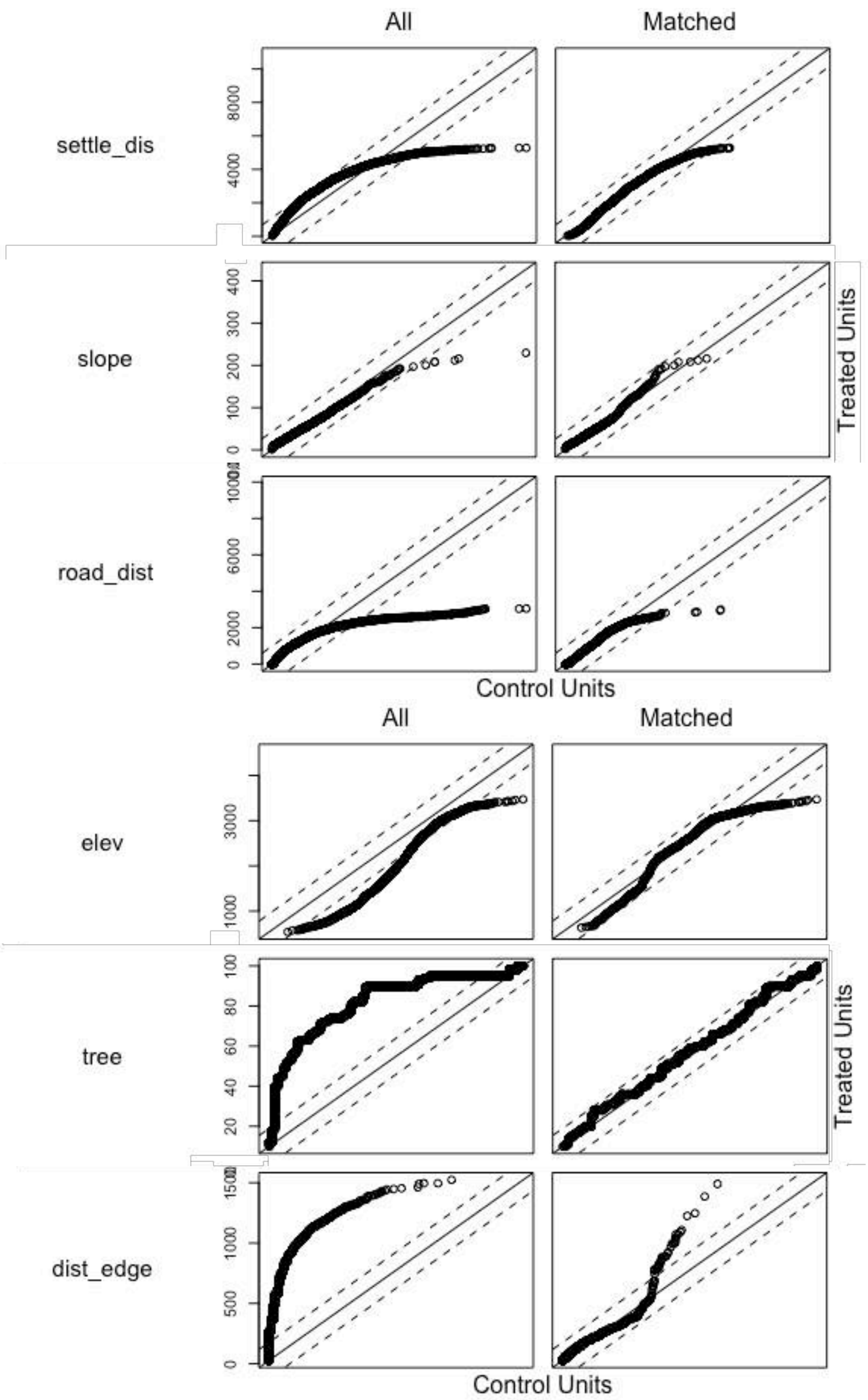

Figure 22. QQ plots of the covariate distributions for Bosques Nublados de Udima (Per) before and after matching. Matched samples N=2834.

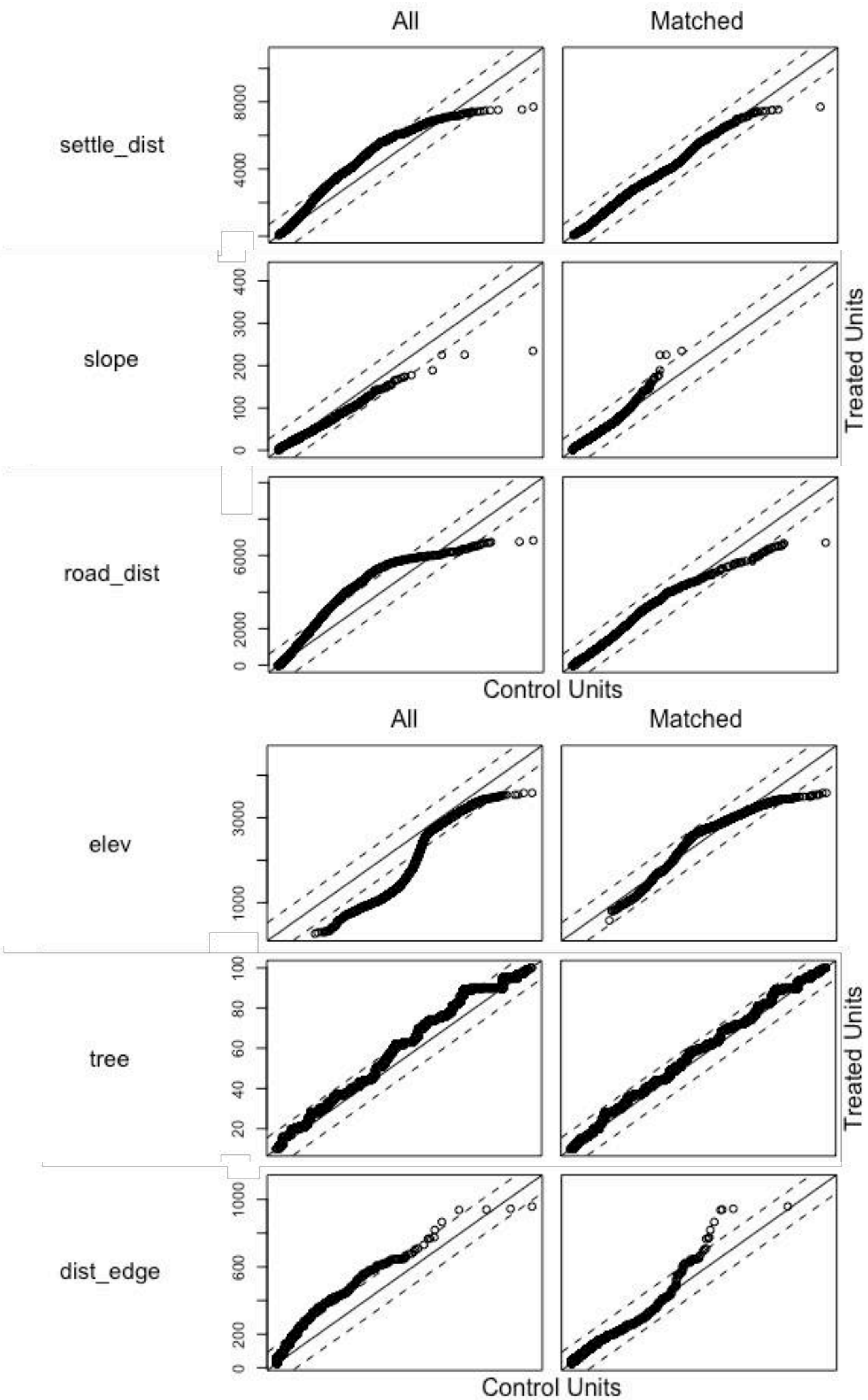

Figure 23. QQ plots of the covariate distributions for Bosques Nublados de Udima (Per) buffer before and after matching. Matched samples N=3956.

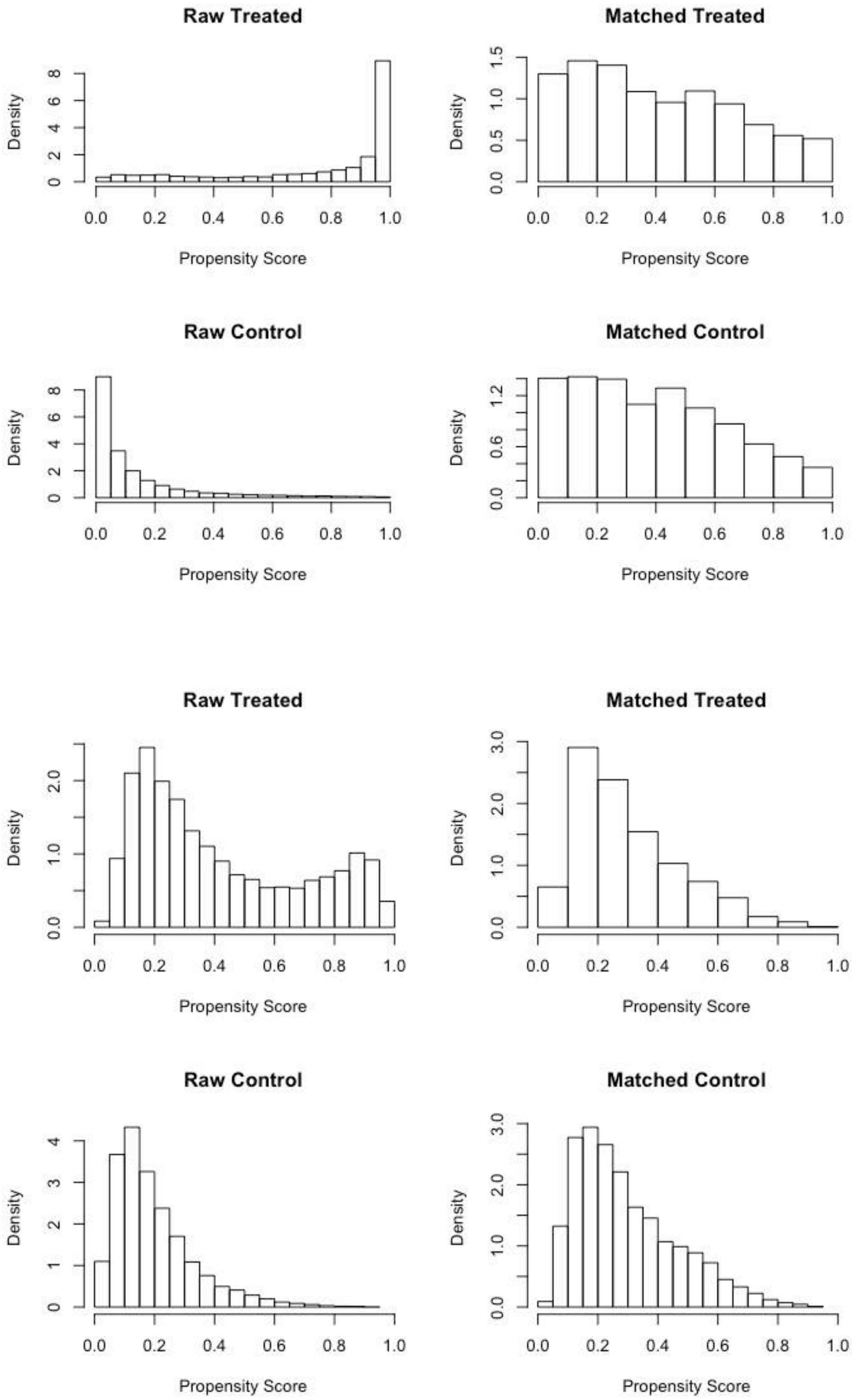

Figure 24. Histogram distributions of Bosques Nublados de Udima (Per) before and after matching for PA (top) and Buffer (bottom).

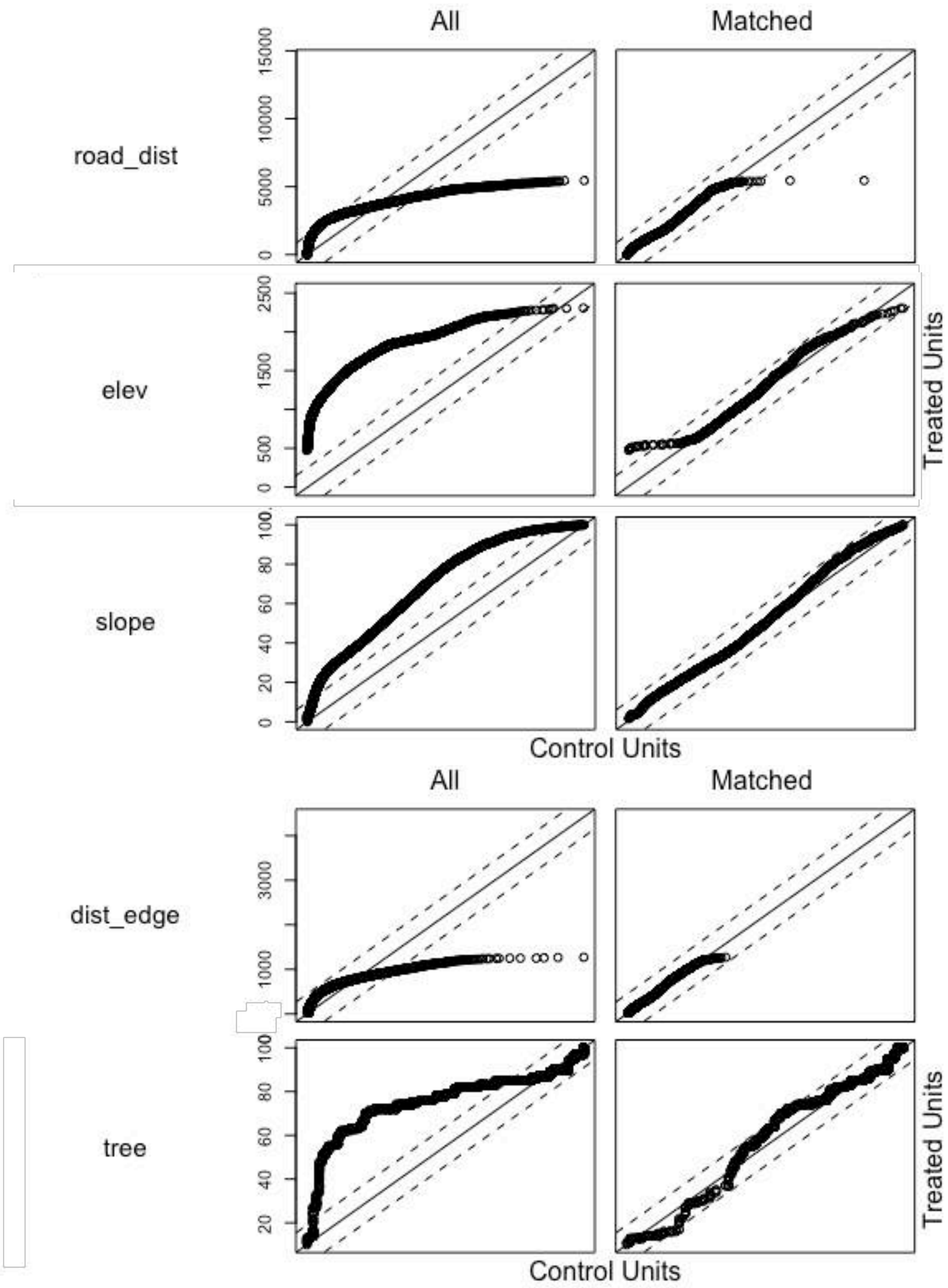

Figure 25. QQ plots of the covariate distributions for Mount Balatukan Range (Phi) before and after matching. Note that due to the low number of settlements on the island distance to settlement was not included. Matched samples N=3794

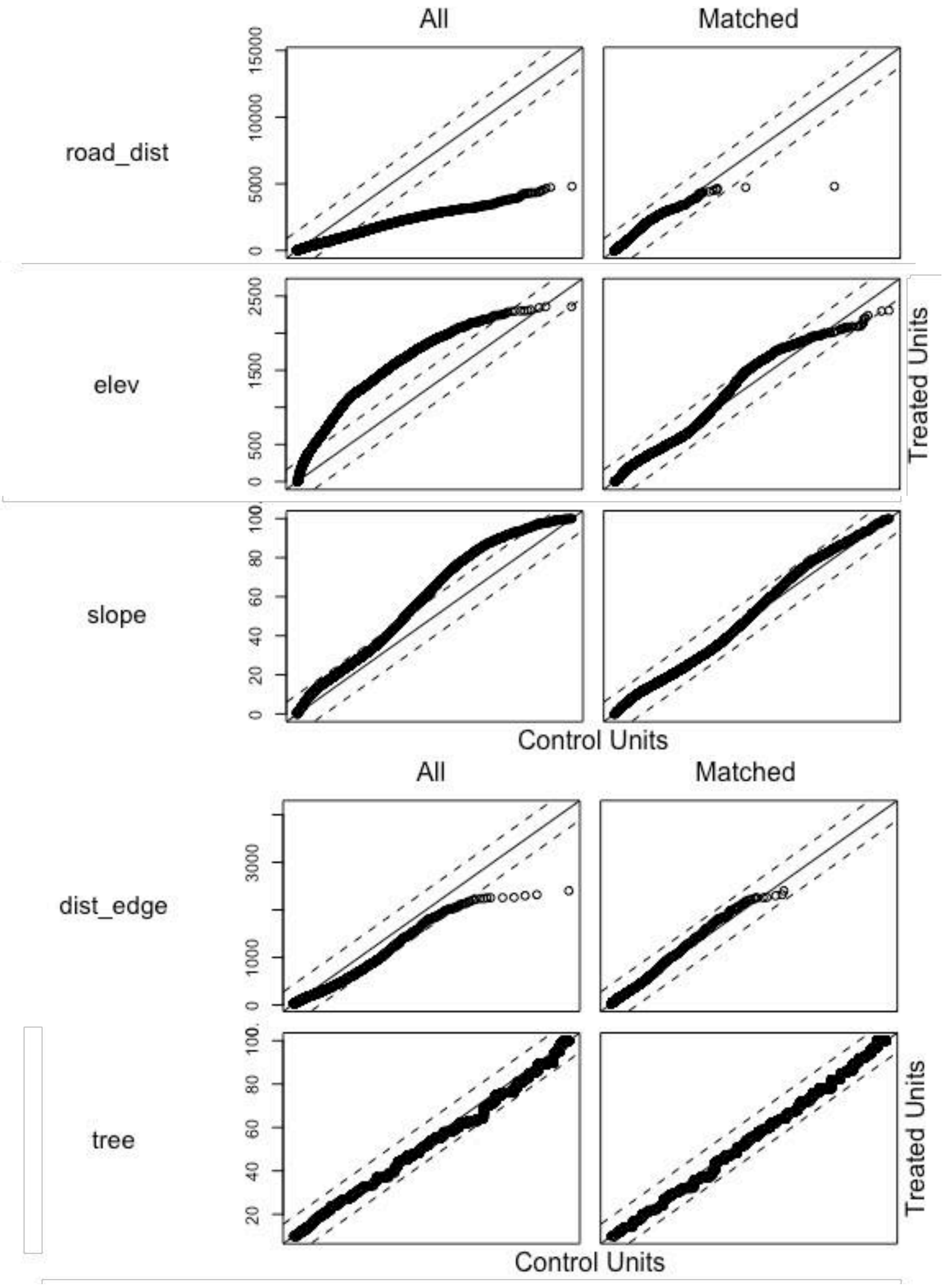

Figure 26. QQ plots of the covariate distributions for Mount Balatukan Range (Phi) buffer before and after matching. Note that due to the low number of settlements on the island distance to settlement was not included. Matched samples N=7033.

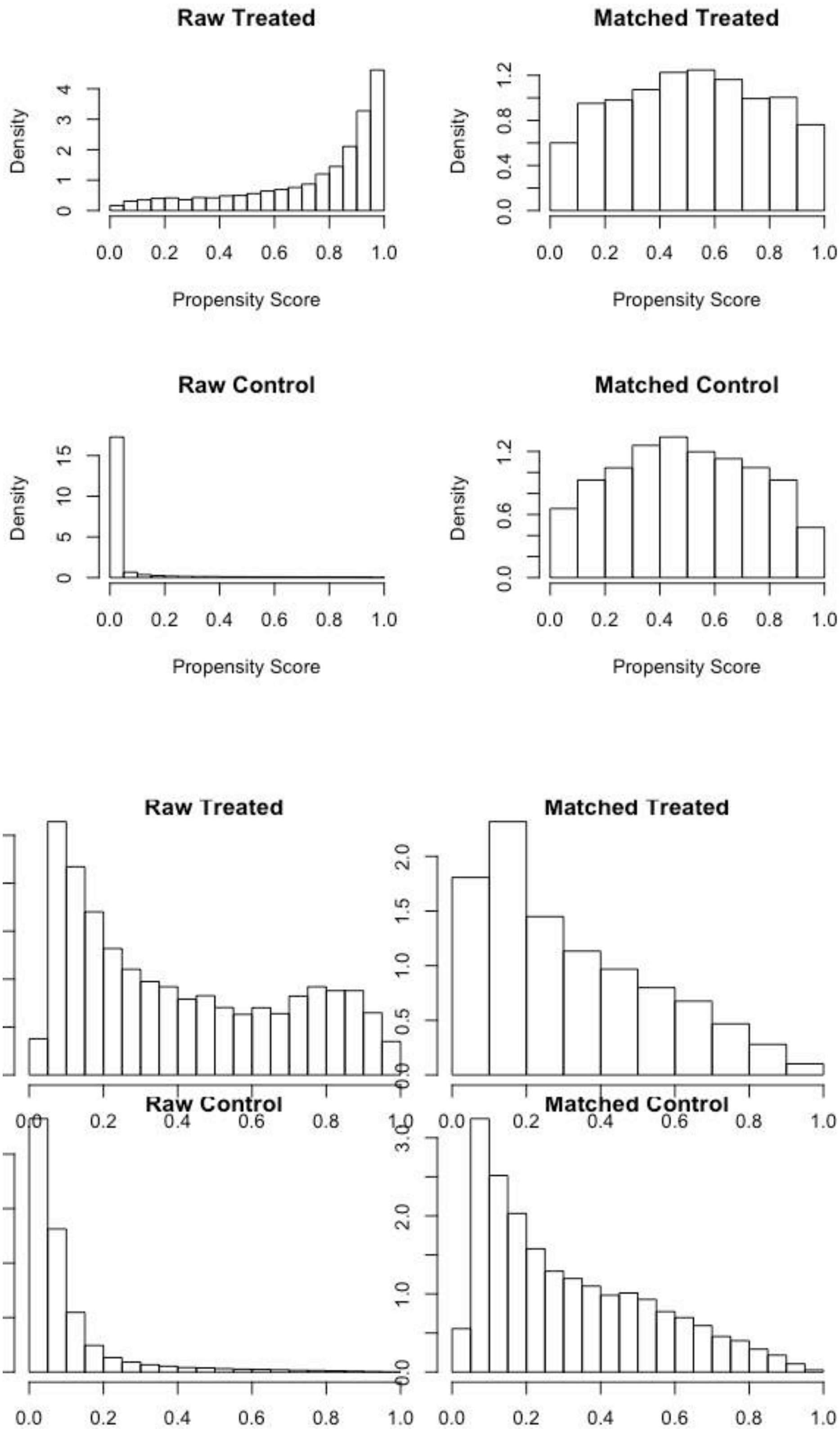

Figure 27. Histogram distributions of Mount Balatukan Range (Phi) before and after matching for PA (top) and Buffer (bottom).

Appendix F

Table 1. Mean deforestation rates before and after PA designation, derived from the matched samples.

	Mean deforestation rate (% yr^{-1})							
	PA		PA control		Buffer		Buffer control	
Protected Area	Before	After	Before	After	Before	After	Before	After
Boumba Bek/Nki (CAM)	0.014	0.000	0.014	0.050	0.012	0.042	0.023	0.076
Deng Deng (CAM)	0.033	0.350	0.073	0.389	0.095	0.558	0.105	0.682
San Miguel de los Farallones (COL)	0.111	0.084	N/A	N/A	0.179	0.115	0.373	0.303
Congolón, Piedra Parada y Coyocutena (HON)	0.134	0.216	0.503	0.644	0.359	0.253	0.427	0.563
Montaña de Botaderos Carlos Escaleras Mejía (HON)	0.725	2.103	0.555	1.073	0.531	1.221	0.586	1.080
Papikonda (IND)	0.017	0.068	0.139	0.201	0.051	0.163	0.205	0.184
Kyauk Pan Taung (MYA)	0.022	0.906	0.206	0.545	0.109	2.385	0.269	0.839
Bosques Nublados de Udima (PER)	0.342	0.157	0.236	0.271	0.365	0.149	0.180	0.222
Mount Balatukan Range (PHI)	0.000	0.079	0.171	0.330	0.052	0.157	0.216	0.459

Appendix G

Table 1. Results of Mann-Whitney U tests between the different treatment groups of the analysis.

Protected Area	Treatment Significance Test (Mann-Whitney U Test) p-value									
	PA					**PA control**	**Buffer**			**Buffer control**
	Before VS After	Before VS Control	After VS Control	Before VS Buffer	After VS Buffer	Before VS After	Before VS After	Before VS Control	After VS Control	Before VS After
Boumba Bek/Nki (CAM)	0.071	1.000	**0.004**	0.868	**<0.001**	0.399	0.161	0.874	0.004	0.371
Deng Deng (CAM)	**0.003**	**0.005**	0.201	0.104	0.073	**<0.001**	**0.001**	0.067	**0.016**	**0.002**
San Miguel de los Farallones (COL)	0.160	N/A	N/A	0.323	0.051	N/A	0.479	**0.006**	**0.014**	0.772
Congolón, Piedra Parada y Coyocutena (HON)	0.162	**0.003**	**0.004**	**0.002**	0.570	0.902	0.101	0.565	**0.006**	**0.004**
Montaña de Botaderos Carlos Escaleras Mejía (HON)	**0.026**	**0.045**	0.130	**0.025**	0.161	0.063	**0.035**	0.554	**0.008**	**0.004**
Papikonda (IND)	**0.005**	**0.021**	**<0.001**	0.353	**<0.001**	0.305	**0.004**	**0.006**	**0.003**	**0.022**
Kyauk Pan Taung (MYA)	**<0.001**	**<0.001**	0.092	**0.002**	**0.041**	**<0.001**	**<0.001**	**0.003**	**0.016**	**0.004**
Bosques Nublados de Udima (PER)	0.213	0.741	0.596	0.470	0.792	0.739	**0.028**	0.056	0.056	**0.004**
Mount Balatukan Range (PHI)	**0.037**	**0.003**	**0.003**	**0.010**	0.056	0.398	0.053	**0.013**	**0.002**	**0.036**

Appendix H

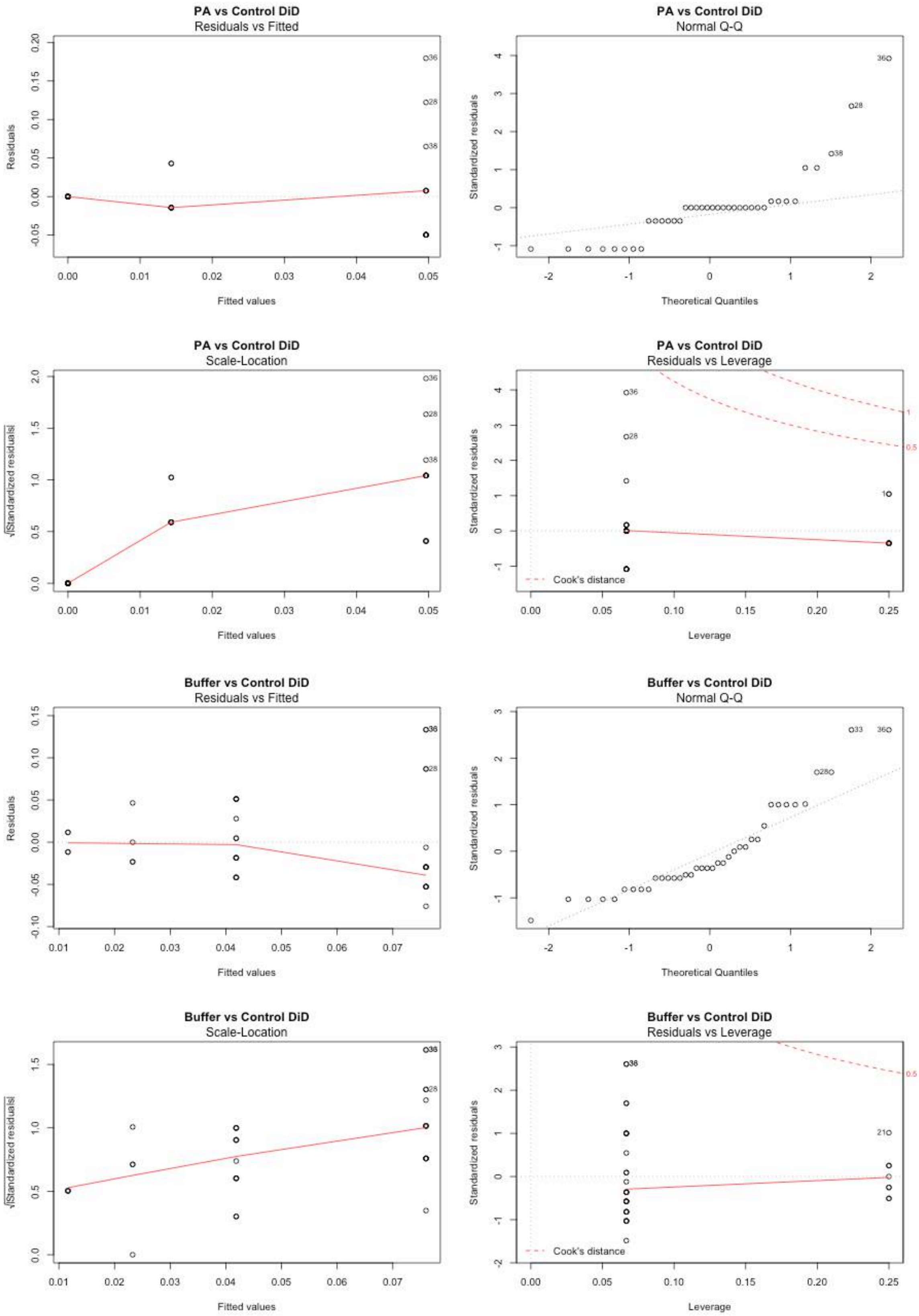

Figure 1. Plots of residuals from the Difference-in-differences linear model for Boumba Bek/Nki (CAM).

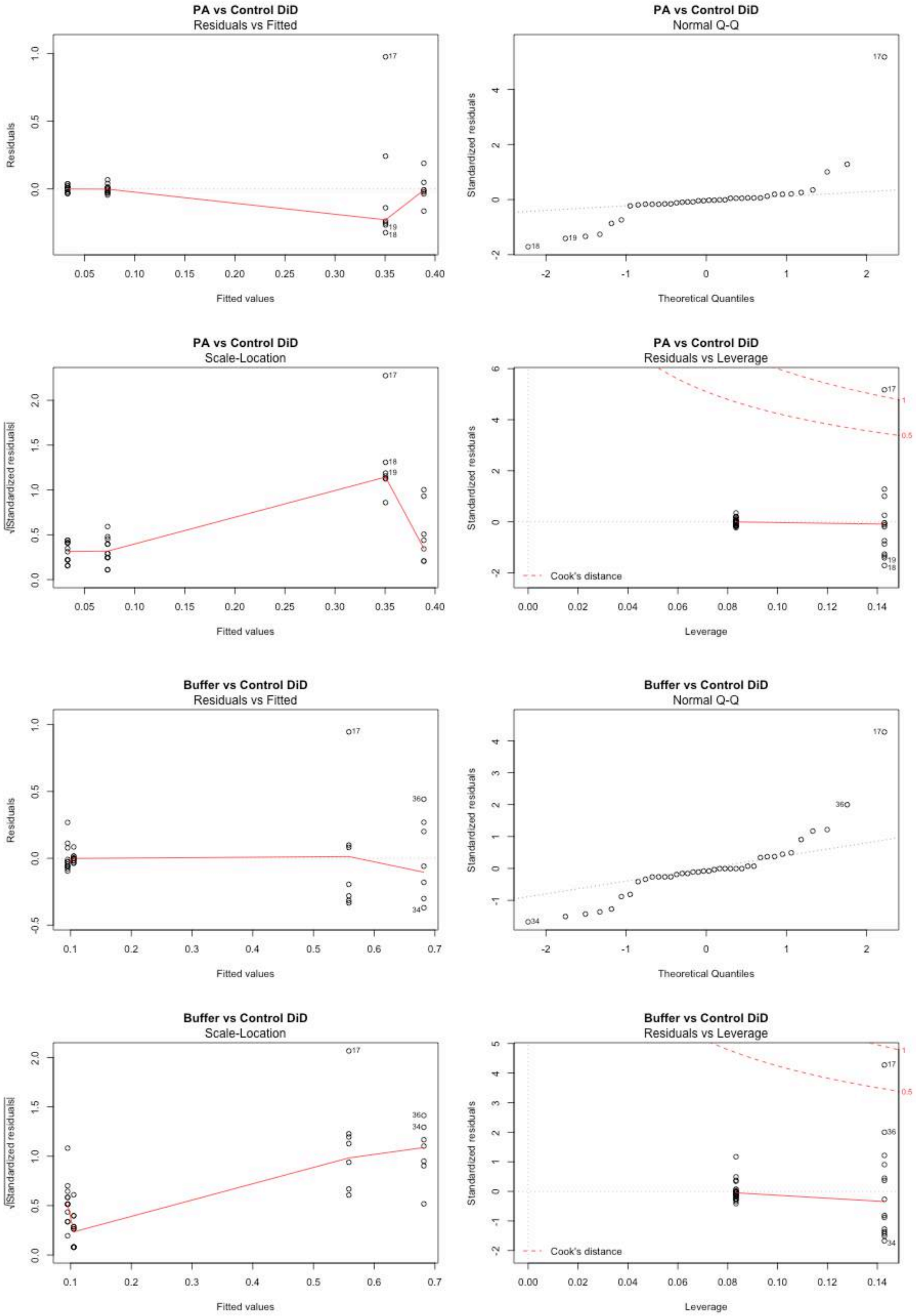

Figure 2. Plots of residuals from the Difference-in-differences linear model for Deng Deng (CAM)

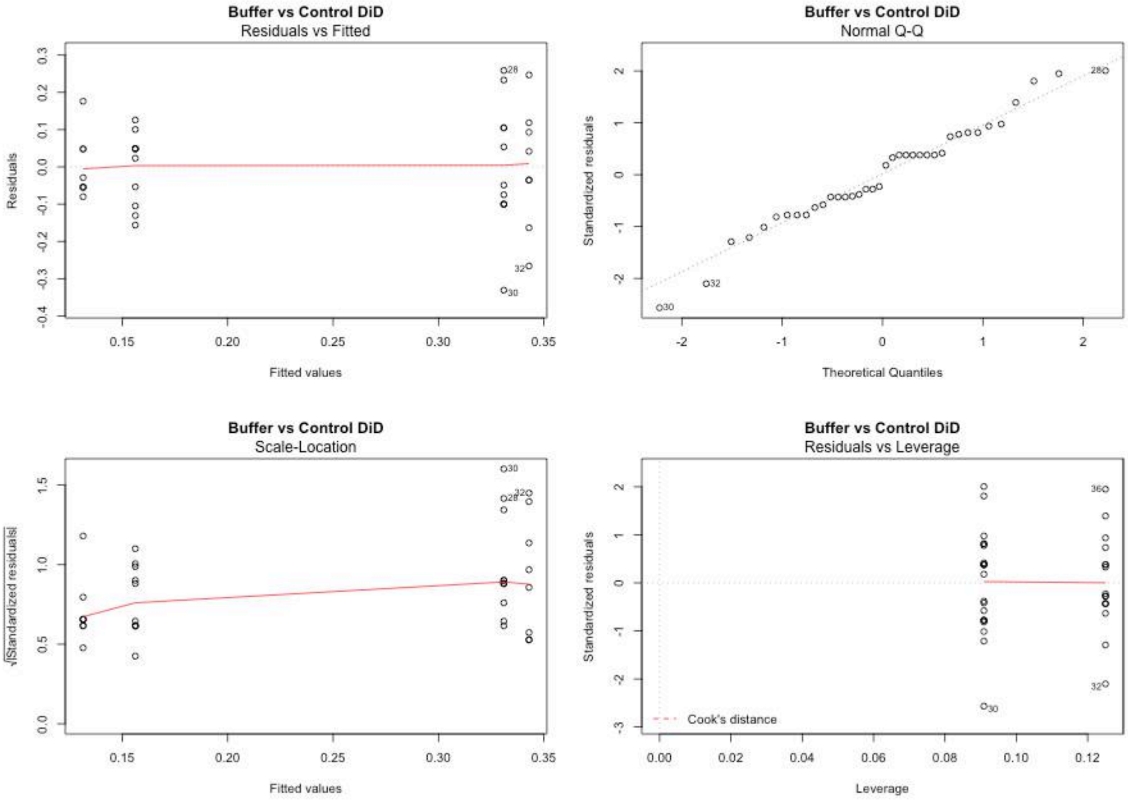

Figure 3. Plots of residuals from the Difference-in-differences linear model for San Miguel de los Farallones (COL).

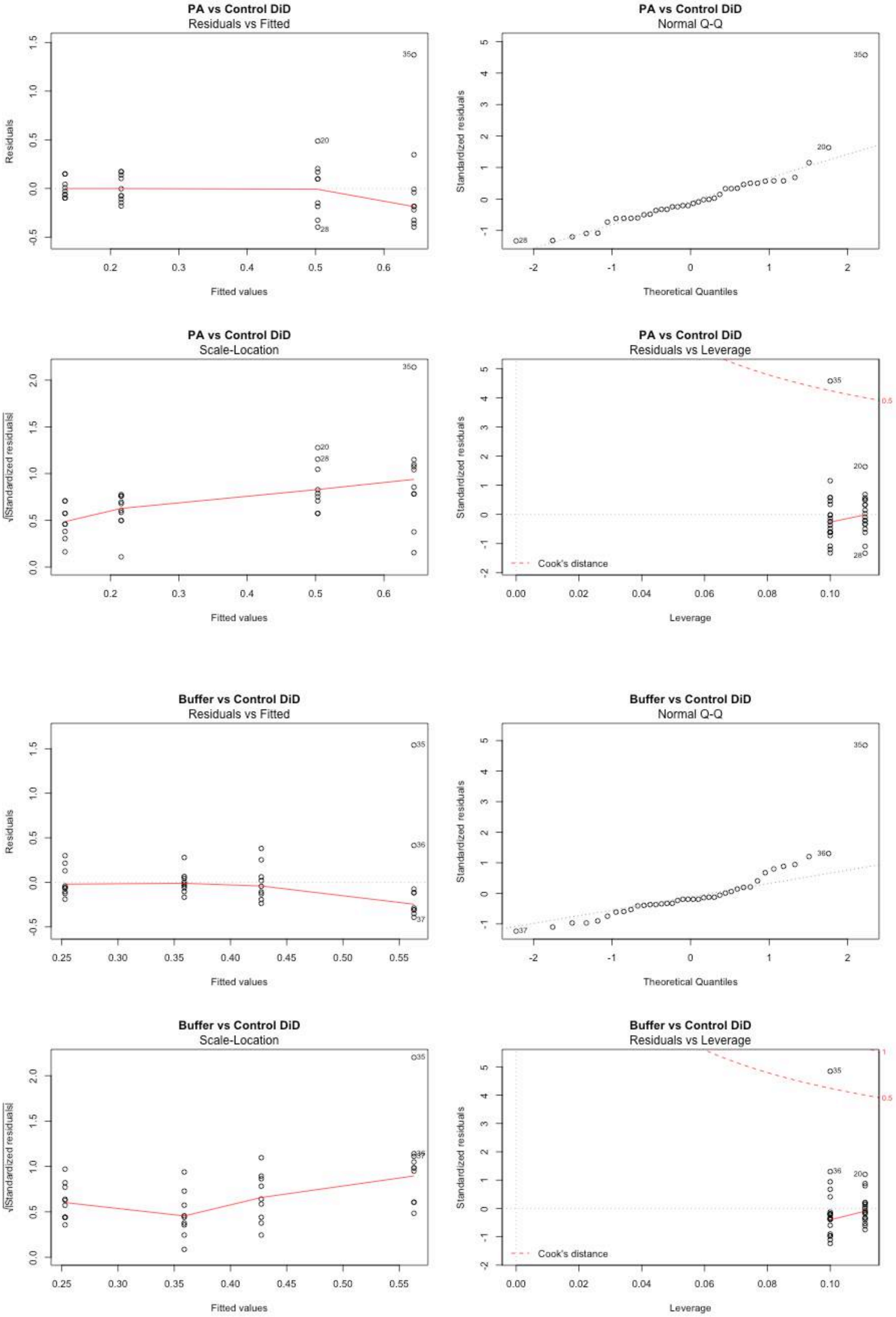

Figure 4. Plots of residuals from the Difference-in-differences linear model for Congolón, Piedra Parada y Coyocutena (HON).

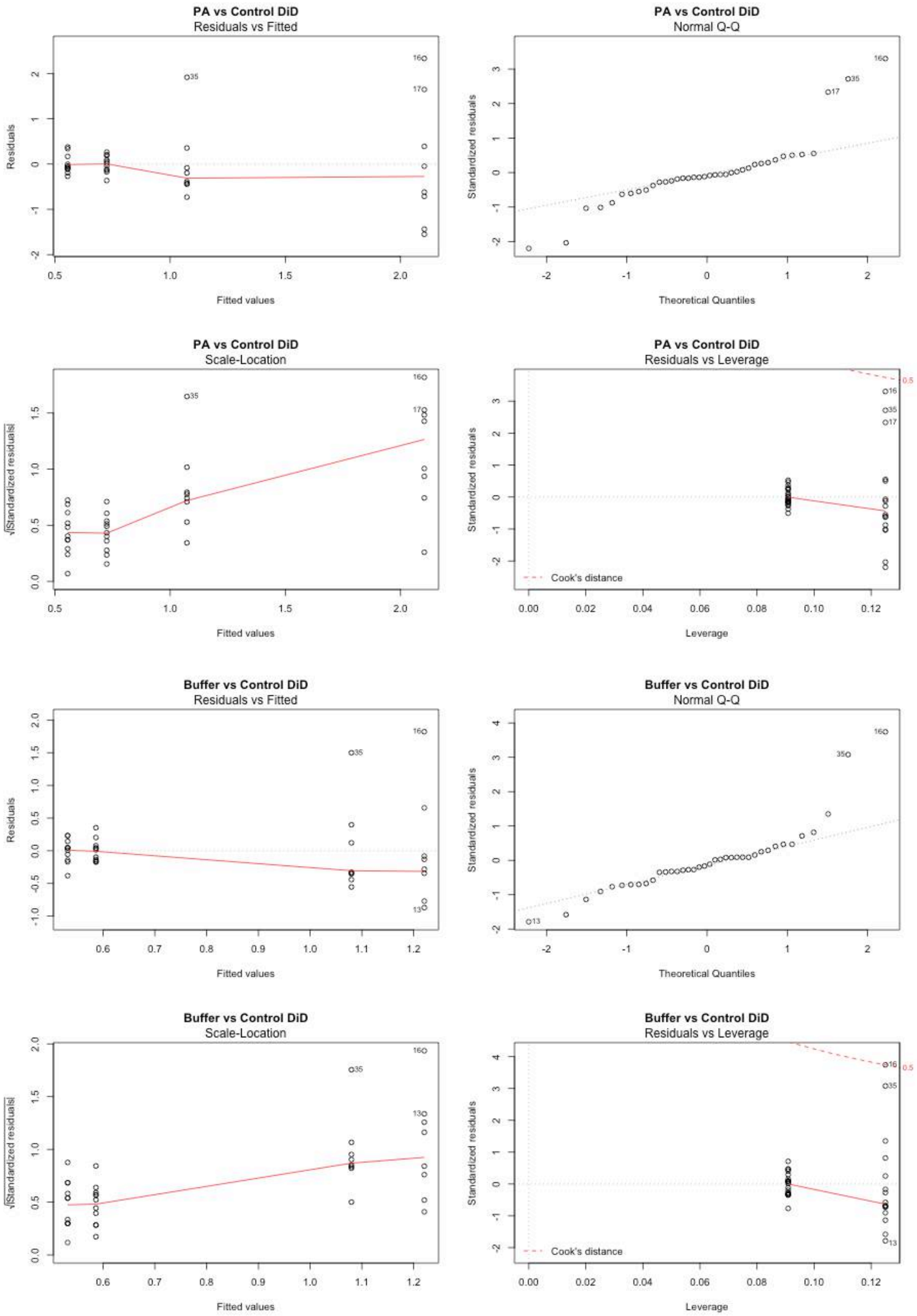

Figure 5. Plots of residuals from the Difference-in-differences linear model for Montaña de Botaderos Carlos Escaleras Mejía (HON).

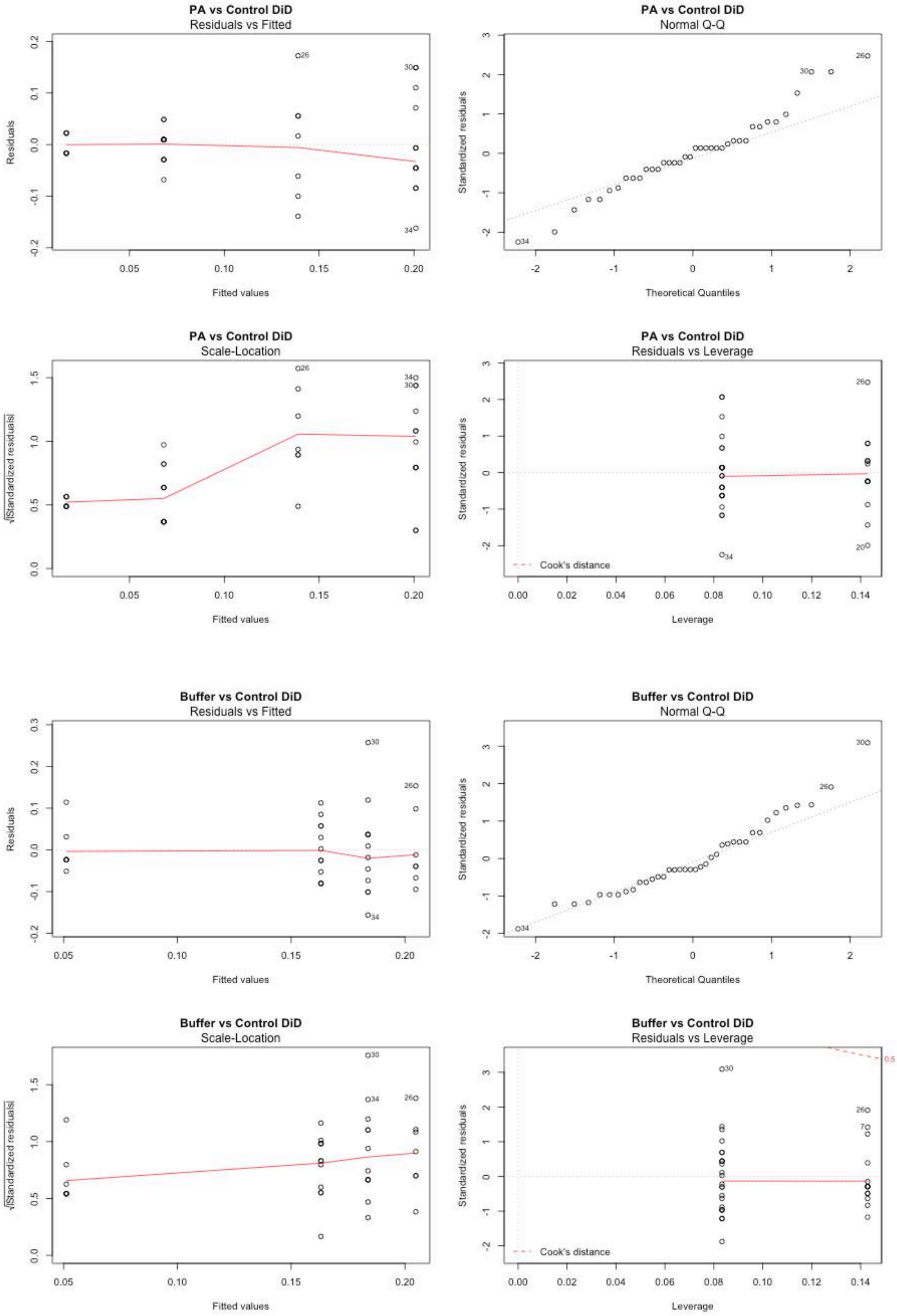

Figure 6. Plots of residuals from the Difference-in-differences linear model for Papikonda (IND).

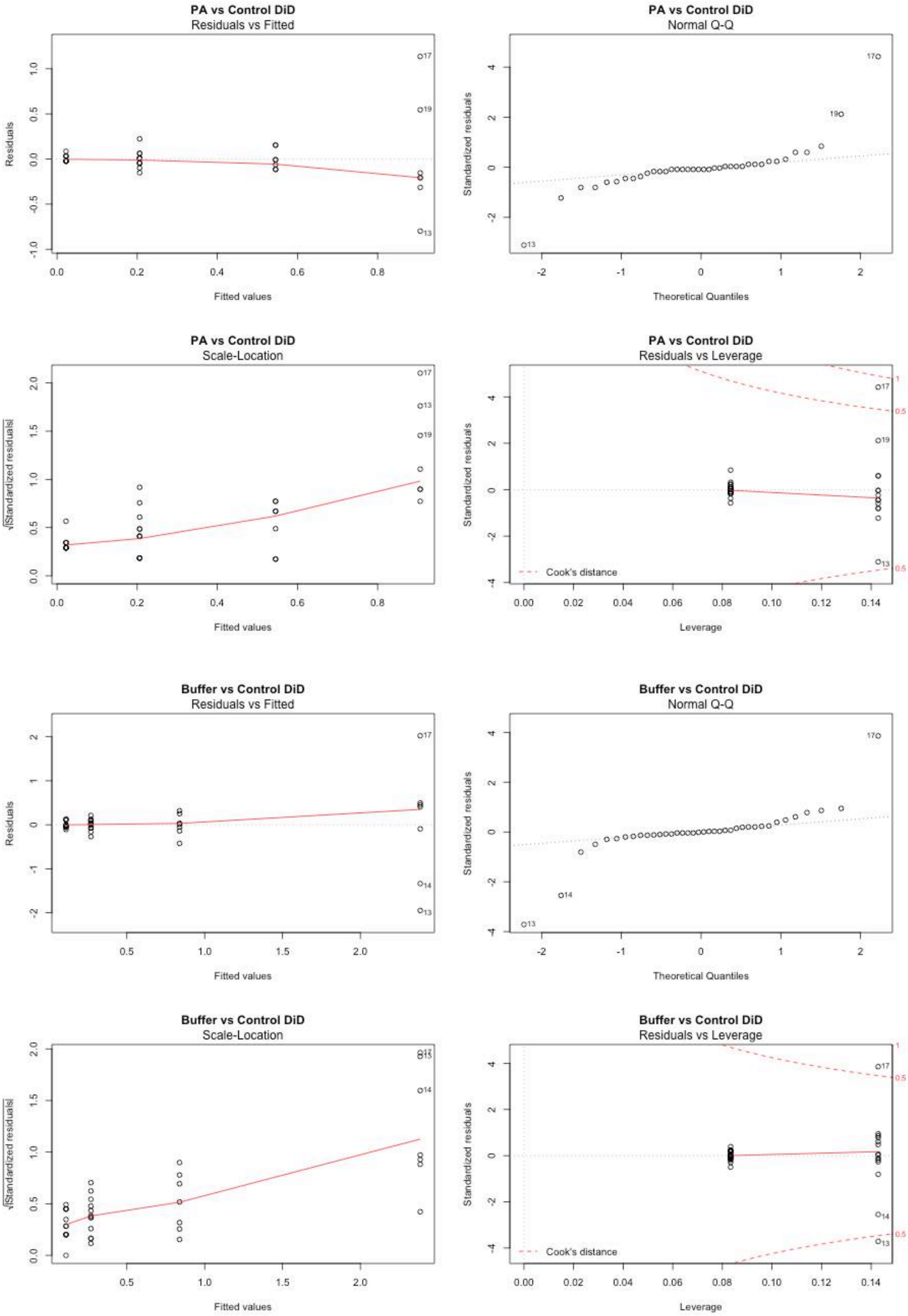

Figure 7. Plots of residuals from the Difference-in-differences linear model for Kyauk Pan Taung (MYA).

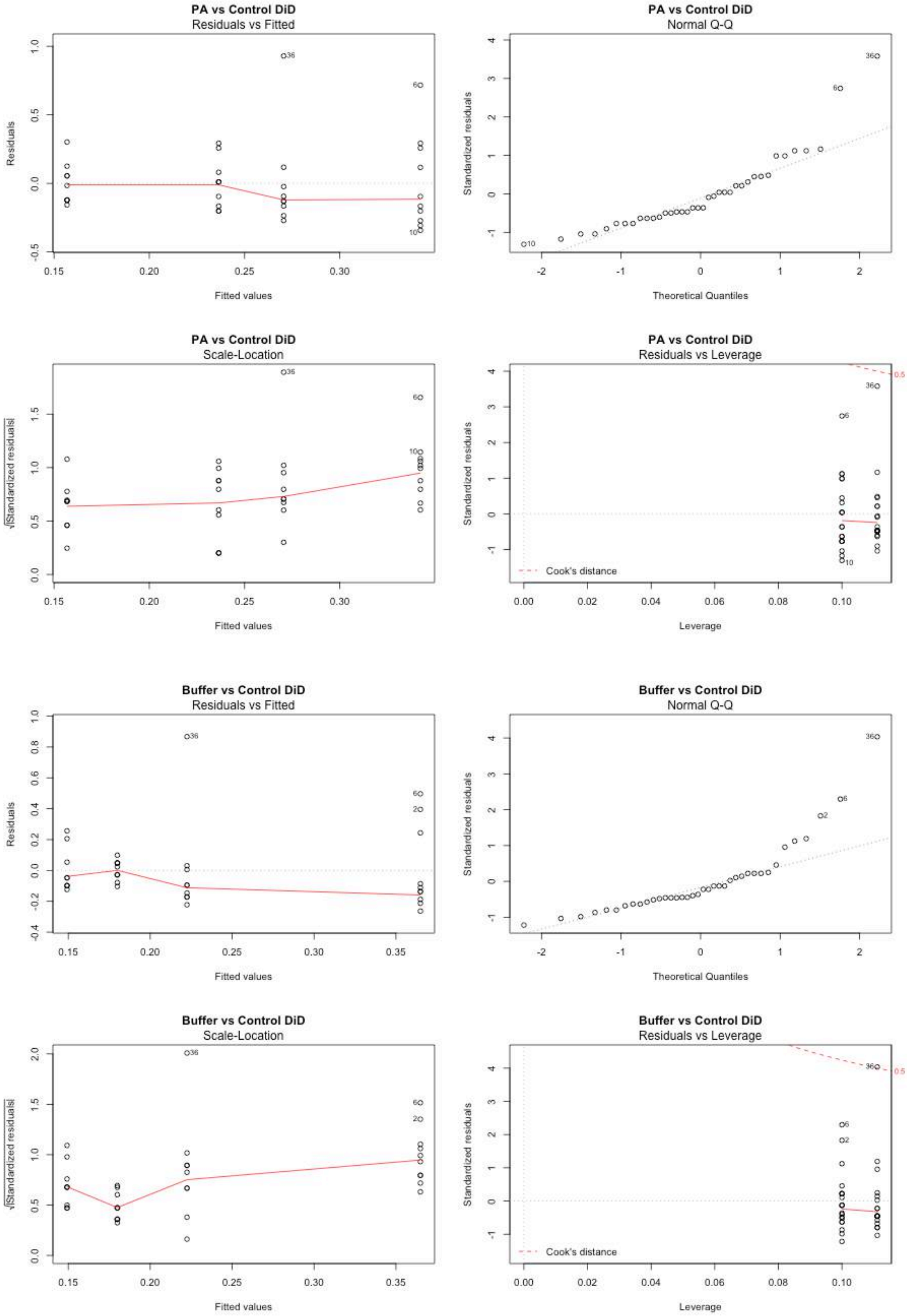

Figure 8. Plots of residuals from the Difference-in-differences linear model for Bosques Nublados de Udima (PER).

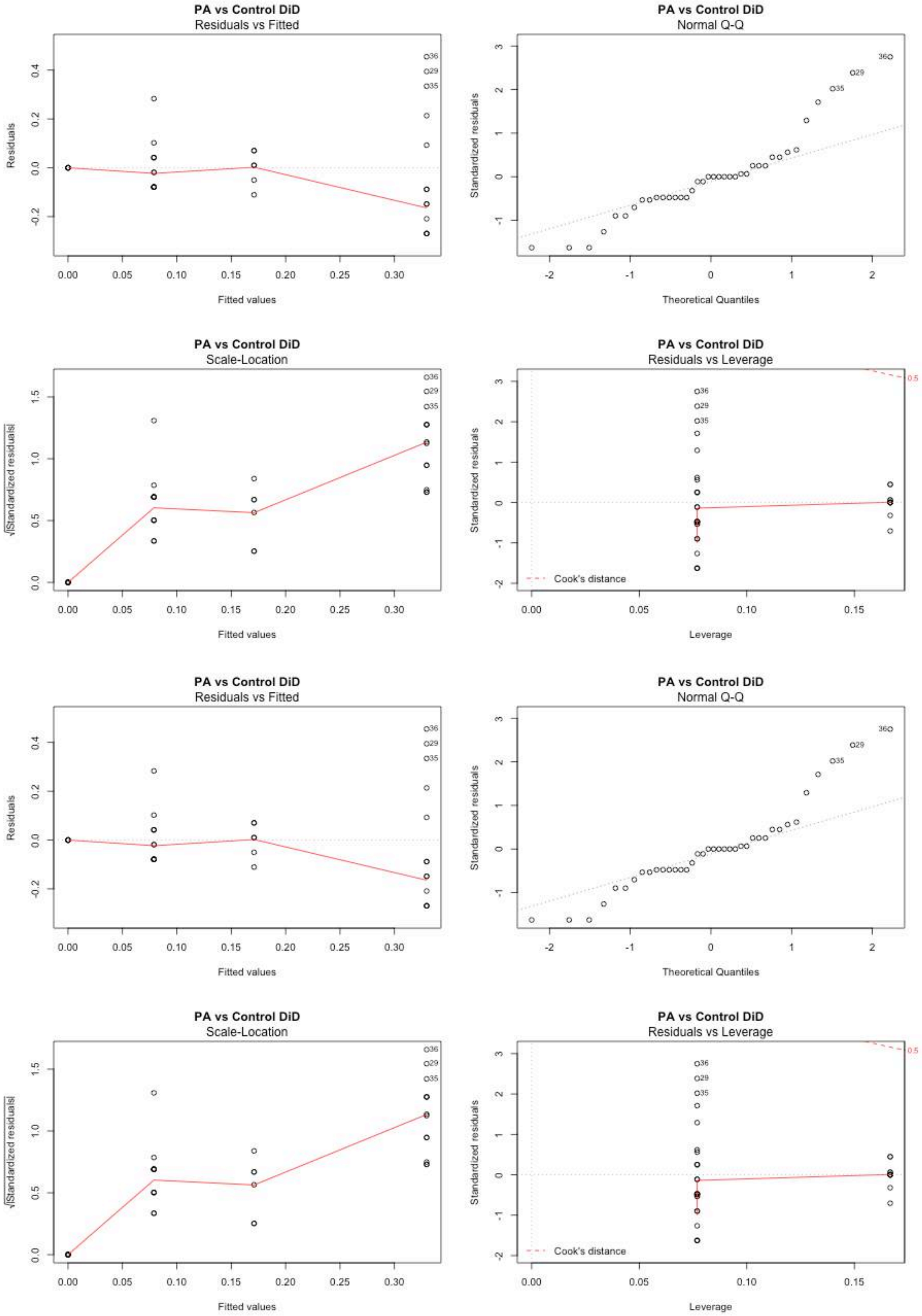

Figure 9. Plots of residuals from the Difference-in-differences linear model for Mount Balatukan Range (PHI).