Sequencing to Satellites:

Integrating biodiversity surveys across scales

Andrew Christopher Slater (B.Sc., M.Sc.)

Lancaster Environment Centre,

Lancaster University

A thesis submitted to Lancaster University in fulfilment of the requirements for the degree of Doctor of Philosophy

April 2025

Declaration

I declare that the work presented in this thesis is my original work and has not been submitted elsewhere, in whole or in part, for any other award or higher degree.

In all sections of this thesis which involve joint research, the collaborators have been acknowledged.

Andrew Slater
Lancaster University
April 2025

Statement of contributions

This thesis has been prepared as chapters in the format of papers intended for submission to peer reviewed journals. With the exception of the General Introduction and General Discussion chapters, each chapter is the product of collaborative work and data sharing. In all cases, A. Slater obtained the relevant satellite imagery, designed the analysis and wrote the chapter using data supplied as detailed below; A. Bush and I. R. Hartley (Lancaster Environment Centre, Lancaster University , Lancaster, UK) provided supervision, helped with development of research questions, data analysis and commented on drafts.

Chapter 2. *Advancing monitoring of a biodiversity hotspot: predicting change in the Succulent Karoo using satellite remote sensing.*
Data supplied by L. Vorster (Sanbona Nature & Wilderness Reserve NPC, Republic of South Africa), N. Saayman (Western Cape Department of Agriculture, Republic of South Africa) and H. van der Merwe (South African Environmental Observation Network (SAEON), Republic of South Africa).

Chapter 3. *Predicting ecological variation in a global biodiversity hotspot: linking metabarcoding to Earth observation.*
Data supplied by D. W. Yu (School of Biological Sciences, University of East Anglia, UK and the Yunnan Key Laboratory of Biodiversity and Ecological Security of Gaoligong Mountain, Kunming Institute of Zoology, Chinese Academy of Sciences, China), and F. Sanderson and F. Edwards (Centre for Conservation Science, Cambridge University, Cambridge, UK).

Chapter 4. *Assessing the biodiversity of bird communities in the Peruvian Amazon is improved by the use of Landsat derived variables.*
Data supplied by C. Kirkby and C. Ketola (Fauna Forever, Peru).

Chapter 5. *Linking fire, logging and bird communities with Earth observation in the Brazilian Amazon.*

Data supplied by O. Metcalfe (Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK).

Acknowledgements

I would like to express my deepest gratitude to Dr. Alex Bush and Dr. Ian Hartley for their continued support, suggestions, guidance and feedback throughout the course of this PhD.

I am grateful to all those people who helped in the collection of field data, and wish to extend my sincere thanks to Liesl Vorster, Dr. Helga van der Merwe, Dr. Nelmarie Saayman, Ollie Metcalfe, Chris Ketola, Dr. Chris Kirkby, and Dr. Doug Yu for agreeing to share their data and taking the time them explain it to me.

Lastly, I would like to mention those that were part of my journey. Lani Edwards for getting me started, Johan Fourie for teaching and instilling a passion for nature, Paul Vorster for giving me the chance to become involved in conservation, and to Dr. Aldina Franco and Prof. Gus Mills for encouraging me to undertake a PhD.

Without these people, this thesis could not have come to fruition.

Abstract

The global decline of natural ecosystems is a consequence of a number of phenomena, including climate change, habitat loss and fragmentation, and the introduction of invasive alien species. As a result, the ecosystem services provided, including carbon sequestration, nutrient cycling through the trophic pyramid, and pollination and seed dispersal, are at risk. Mitigation strategies to halt and reverse these trends can be devised but must be informed by spatially and temporally accurate information on the location and causes of ecosystem change. This type of scalable ecological modelling directly supports global biodiversity frameworks, such as the Kunming-Montreal Global Biodiversity Framework (GBF), which call for spatially explicit monitoring of ecosystem condition and change.

While field surveys can provide accurate information to inform mitigation strategies for target ecosystems, they are often costly in terms of time and resources. Consequently, the frequency and spatial and temporal extent at which field surveys can be conducted is limited. In contrast, Earth observation (EO) satellites provide freely available, repeated, wall-to-wall information of the Earth’s surface. The combination of EO data with environmental data has been successfully employed to describe land classes, and quantify changes within them, such as levels of deforestation. Although the description of land classes enables some inference of expected ecosystem structure, ecosystems do not always transition categorically, although abrupt land use change can result in clear shifts. More often, ecosystems vary gradually across landscapes, much like the reflectance values recorded by EO satellites. It is hypothesised that the composition of the studied communities will vary in accordance with the structure of the habitat, and that elements of this structure will be described by EO.

The objective of this thesis was to investigate the potential of combining EO data with field survey data in multi-species distribution and occupancy models to predict community composition across landscapes. This approach enables the investigation of the influence of environmental factors on biodiversity and the inference of ecosystem condition. The study used four distinct groups of field data: plant communities in the Succulent Karoo of South Africa, insect communities in the vicinity of the Gola Rainforest in Sierra Leone, and bird communities identified in the Peruvian and north Brazilian Amazon. These data sets comprised between 120 and 235 species, which were included in analysis.

The study found that for all data sets, models fitted with EO data yielded predictions of community composition that were as or more accurate than those made by models fitted with environmental data. In predicting the validation data, the community mean area under the curve (AUC) values ranged between 0.58 and 0.66. The models generated AUC's exceeding 0.7 for between 21% and 49% of the individual species modelled. These values represent a moderate predictive performance overall, with the accuracy of models influenced by the strength and extent of ecological gradients, species detectability, and survey design.

The findings indicate that there are consistent associations between EO variables and a large number of species from diverse taxonomic groups and geographical locations. It can be reasonably assumed that a significantly higher proportion of taxa could be predicted using similar methodologies provided that sufficient field surveys are conducted to calibrate similar models. This highlights the importance of obtaining high-quality ecological data through field surveys in order to calibrate our interpretation of ecological changes observed through remote sensing. The ability to apply EO data retrospectively and prospectively also offers a valuable opportunity for long-term biodiversity monitoring across changing landscapes.

Table of Contents

List of Figures

List of Tables

1. General Introduction

1.1 Biodiversity declines and the need for scalable monitoring

In the face of an alarming decline in global taxonomic and functional biodiversity, environmental issues have become a pressing concern of the 21st century (Jenkins and Joppa, 2009; Laurance et al., 2012; Pimm et al., 2014). As a well-studied group, the extinction risks and rates faced by many vertebrate species are well documented. Over the past two centuries, the extinction rate of bird species has been estimated to be up to 100 times higher than background levels (Pimm et al., 2006). While tens of mammal species have gone extinct in recent decades, nearly 200 (4%) of the world's mammal species have lost more than 50% of their range since the 19th century (Ceballos and Ehrlich, 2002). The decline of mammal populations may be a more informative indicator of environmental condition than species extinctions alone. Insects outnumber vertebrates by orders of magnitude, and while only tens of insect species have been officially recorded as becoming extinct in modern times, the actual number is likely to be in the thousands (Dunn, 2005). Over the past century, plant extinctions have occurred at nearly 2.5 species per year, more than twice the rate observed before 1900 (Humphreys et al., 2019). Certain ecosystems play a disproportionate role in sustaining global biodiversity. Tropical forests and the Greater Cape Floristic Region of South Africa, for example, support over half of the world's biodiversity, yet cover less than 10% of the Earth's surface, making them key conservation priorities (Bradshaw et al., 2009; Gibson et al., 2011; Slingsby et al., 2017). These losses are largely driven by climate change, invasive species, and habitat loss, all of which are exacerbated by human activities.

Conservation actions play a crucial role in reducing biodiversity loss (Langhammer et al., 2024). Without intervention, bird and mammal extinctions in recent decades could have been three times higher (Bolam et al., 2021). To effectively mitigate anthropogenic impacts on biodiversity, conservation strategies must be guided by current evidence, adaptable to local conditions, and be informed by the successes and failures of past actions (Pettorelli et al., 2014b). Effective biodiversity monitoring, in terms of measuring species richness and composition, is essential for assessing ecosystem health and guiding conservation strategies (GEO BON, 2018; Pereira et al., 2013). Environmental condition can be described structurally in terms of type (e.g. forest, agriculture), quantity (e.g. percentage of each) and distribution of elements within it (Pettorelli and Schulte to Bühne, 2022). Alternatively, environmental condition can be inferred from the richness and diversity of faunal indicator taxa. For example, bird community richness is used to evaluate the effectiveness of different reforestation techniques (Barros et al., 2022); while insect diversity serves as a bioindicator of terrestrial, airborne, and aquatic pollution levels (Parikh, Rawtani and Khatri, 2021).

Traditional field surveys, while essential for biodiversity monitoring, are constrained by financial limitations, accessibility issues and scalability challenges (Rocchini et al., 2016; Skidmore et al., 2015). Satellite remote sensing has emerged as a pivotal tool, offering a cost-effective way of monitoring land cover and habitat change over large geographic areas. The Landsat programme started in 1972 and provides medium-resolution multi-spectral imagery suitable for long-term vegetation monitoring (Wulder et al., 2019). The Sentinel-2 satellites, part of the European Space Agency's Copernicus Programme, deliver higher-resolution optical data with 10-60m pixel sizes, with a 5-day revisit time between them, compared to an 8-day revisit from combined Landsat satellites (Drusch et al., 2012; Wulder et al., 2019). MODIS (Moderate Resolution Imaging Spectroradiometer) has offered near-daily global coverage since 1999 and is particularly useful for tracking vegetation changes at a coarser 500-1000m resolution (Justice et al., 2002). The Sentinel-1 mission is a sister programme to Sentinel-2 and provides radar data which is unaffected by clouds and independent of daylight, enabling surface structure and moisture analysis (Torres et al., 2012). NASA's GEDI (Global Ecosystem Dynamics Investigation) is a spaceborne LiDAR instrument that captured vertical forest structure in pulses with 30m footprints between 2019-2022 (Dubayah et al., 2020). While most datasets from these missions are freely available, PlanetScope is a commercial satellite constellation offering daily imagery at ~4m resolution, which has provided data to academic researchers through research partnerships.

Variations in plant community structure and underlying substrates affect spectral reflectance that can be detected by satellite sensors (Frye et al., 2021; McDowell et al., 2015). Vegetation structure influences faunal communities and the functional biodiversity of an area (Stein et al., 2014), making changes in reflectance observed by satellites useful indicators of ecosystem dynamics. Satellite imagery can categorise broad landscape types, and infer finer variations associated with species richness, distribution, abundance, and community composition (Skidmore et al., 2021). While Earth observation is valuable for capturing ecosystem dynamics (Regos et al., 2022; Wulder et al., 2016), it cannot directly measure biodiversity in its entirety (Bush et al., 2017). Thus, there is a trade-off between the detail of what can be measured *in situ* and what is observable from satellite imagery. Despite this limitation, Earth observation data have been successfully used in classifying land cover and tracking areas and times of deforestation (Hansen et al., 2013) or land class (Brown et al., 2022). Earth observation has also been used to map dominant invasive plants in grasslands (Xing et al., 2021), and assess the distribution and richness of large, observable species such as oak (*Quercus* spp.) trees (Pinto-Ledezma and Cavender-Bares, 2021). To examine how biodiversity monitoring challenges vary across ecosystems and taxa, this study considers variation in species detectability, the effectiveness of remote sensing techniques and ecological responses to environmental variables. It applies these methods in diverse environments, including tropical forests and semi-arid regions, and focuses on a range of taxa such as birds, insects and plants.

1.2 Gaps in remote sensing for biodiversity monitoring

While remote sensing provides a valuable tool for large-scale biodiversity assessment, challenges remain in how ecological communities are represented and analysed using satellite data. Describing a landscape by land class has utility, and community structure can be influenced by a single dominant species. However, describing the community as a whole is sometimes a more important biodiversity metric (Skidmore et al., 2021). The definition and selection of classes with which to analyse a landscape can be prone to researcher bias and require prior knowledge. But continuous descriptors of the land’s surface, such as those provided by satellite imagery, can classify landscapes at a finer resolution than is traditionally achievable by field ecologists (Nguyen Trong et al., 2020; Pettorelli et al., 2014b). More recently, studies suggest that the heterogeneity of pixel values in satellite imagery can be used to infer ecological niches on the ground, and surface reflectance has been shown to better describe community composition and richness compared to categorical variables (Perrone et al., 2023; Purdon et al., 2022).

The growing emphasis on understanding biodiversity patterns at the community level, rather than focusing on individual species, has led to a growing trend towards using joint species modelling, particularly in the context of ecological monitoring and conservation planning (Ferrarini et al., 2021). Traditional distribution models infer correlations between the presence or absence of a species and environmental covariates. These correlations are then used to map species distributions across a landscape based on known covariate values. To predict communities, stacked distribution models overlay the outputs of multiple independently predicted species. However, these models do not account for interactions between species and traditionally overestimate species richness (Kissling et al., 2012). Joint species distribution models simultaneously model multiple species while accounting for residual correlations that remain after environmental covariates have been fitted (Wilkinson et al., 2021). Occupancy models take distribution models a step further by using results from multiple surveys and allowing for imperfect detection when inferring relationships between a species' distribution and environmental covariates. Multi-species occupancy models are extensions of joint species distribution models, but while they account for imperfect detection, until recently they have rarely accounted for residual autocorrelation between species (Doser et al., 2023). Continual advances in statistical modelling are being made, with many recent models now incorporating spatial autocorrelation, and some allowing for the inclusion of species traits or phylogenies (Ovaskainen et al., 2017). These models can be used to link satellite remote sensing data with biodiversity survey data to estimate entire communities continuously across a landscape (Csillik et al., 2019). Despite the increasing number of studies linking remote sensing to ecological patterns, understanding how well these metrics actually represent ecological communities remains a key challenge. Without a clearer understanding of these relationships, remote sensing is at risk of being a tool for classification rather than for true biodiversity assessment (Regos et al., 2022).

1.3 Research aims questions and hypotheses

There are ongoing challenges in linking remote sensing data with ecological survey results, particularly in capturing community structure, and spatial patterns using satellite-derived variables. This study addresses these challenges by testing whether environmental metrics derived from Earth observation can effectively describe biodiversity patterns across a range of taxa and ecosystems. In addition to recording the presence or absence of target taxa, each biodiversity survey recorded environmental variables at the survey site that were of ecological interest to the study. Variables differed between studies, but included successional stage, proportion of deforestation, evidence of logging or burning, elevation, mean annual temperature, and precipitation.

Question 1: How do *in situ* environmental variables influence the community composition of target taxa? We predict that variables such as habitat structure, climate and disturbance history, will significantly explain variations in community composition and richness across sites.

Question 2: To what extent can environmental metrics derived from Earth observation predict the results of biodiversity surveys? We predict that models incorporating remote sensing variables, such as vegetation indices and land cover structure, will outperform models based solely on *in situ* variables by capturing landscape-scale variation not measured during field surveys.

Question 3: How does satellite platform choice and spatial scale of measurement influence model performance in predicting biodiversity patterns? We predict that finer-scale Earth observation data (for example higher resolution imagery with more frequent revisits) will yield stronger predictions than coarser scale data. However, the optimal scale may vary between ecosystems and taxonomic groups.

Question 4: Can Earth observation data be used to generate reliable biodiversity maps, and how remote sensing might be used to improve the efficiency of ecological surveys? We predict that biodiversity maps generated using satellite-derived variables will align closely with field data and provide a cost-effective way of extrapolating biodiversity patterns across unsampled areas. We further predict that these maps will offer a more nuanced representation of community composition, and of how and where biodiversity changes across the landscape, than traditional land class, biome, or ecosystem type classifications.

1.4 Thesis structure

This thesis is divided into six (6) chapters. The current chapter (Chapter **1**) provides a general context for the study. The objectives of the research are identified, and their practical value is presented.

Chapter 2 examines plant communities in the Greater Cape Floristic Region of South Africa. Plant communities are predicted to directly influence surface reflectance values, and I assess the extent to which optical satellite data, in combination with elevation, temperature, and slope aspect can describe plant community composition and abundance. Communities are mapped based on their similarity in composition, and areas of directional change are identified. This chapter primarily addresses Questions 1, 2 and 4.

Chapter 3 examines insect communities in the community forest surrounding the Gola Rainforest National Park, a REDD+ site in Sierra Leone. Insect communities were assessed using DNA metabarcoding of trapped samples, with communities modelled against satellite data and *in situ* descriptors of forest structure. The study aims to identify areas where cocoa farming practices could be intensified while minimising biodiversity loss. I also investigate how different satellite platforms and spatial scales affect model performance. This chapter contributes to all four questions.

Chapter 4 examines bird communities in the tropical forests of the Madre de Dios region of south-eastern Peru. Multiple surveys were conducted, and occupancy models were used to account for imperfect detection in analysing species distributions. The predictive performance of models using satellite data is compared with those based on landscape type and levels of deforestation. Additionally, I assess how survey design influences the ability to detect changes in these communities and suggest strategies for improving survey efficiency. This chapter primarily addresses Questions 1 and 2.

Chapter 5 examines bird communities identified through passive acoustic monitoring in the tropical forests of Pará in the eastern Brazilian Amazon. Multi-day continuous acoustic recordings were used to create multiple replicate surveys at each site. Surveys were conducted in primary and secondary forests, some of which had been logged and/or burned. Occupancy models were used to explore how habitat conditions influence bird communities. Model predictions using EO data are compared with those based on field-based habitat data. This chapter contributes to Questions 1, 2, and 3.

Chapter 6 provides a synopsis of the thesis, highlighting key findings and suggesting how they may be applied in practice and developed in future research.

2. Advancing monitoring of a biodiversity hotspot: predicting change in the Succulent Karoo using satellite remote sensing

2.1 Abstract

Reliable landscape-scale monitoring is key to informing policy and guiding strategic intervention to halt and reverse global biodiversity loss. Time and resources constrain the scale of field surveys, but Earth observation (EO) satellites provide routine wall-to-wall coverage of the Earth. If changes in vegetation composition or abundance are captured in pixel reflectance values, we can combine restricted field monitoring with EO data to improve our inferences at the landscape scale.

We investigated whether EO data improved our capacity to predict the spatial distribution and temporal dynamics of vegetation in the Greater Cape Floristic Region (GCFR) of South Africa. Our analysis was based on the 211 most frequently observed plant species recorded in 1440 surveys. The value of EO was assessed based on the improvement to joint species distribution models (JSDMs) that were fitted with standard static environmental data. Topography and temperature were the most influential environmental drivers in both distribution and abundance models. The addition of EO resulted in a marginal increase in the explanatory power of distribution models (i.e., presence/absence) by 3%, while a more substantial enhancement was observed in species abundance models, with an increase of up to 30%. Nevertheless, the proportion of variance explained by EO was much greater, representing between 34% and 64% of the total. The inclusion of measurable EO variables replaced much of the residual variance that was otherwise explained by estimated spatial latent variables, allowing for more accurate predictions of composition across the landscape.

We demonstrate that when diverse field data on abundance is available to train models, EO could substantially improve our capacity to identify spatial variation in abundance and temporal changes in the composition of highly diverse communities, such as those found in the GCFR.

2.2 Introduction

The current rate of extinction among plants is unprecedented (Humphreys et al., 2019), and there is evidence that biodiversity and associated environmental services will continue to decline throughout the 21st century (Pimm et al., 2014). The Global Biodiversity Framework (United Nations, 2021) set goals to reduce, halt, and ultimately reverse these trends. That the preceding Aichi Biodiversity Targets (CBD, 2010) were largely missed highlights the fundamental importance of implementing a continuous process of monitoring, measurement, and analysis to assess the efficacy of these targets and our progress towards their fulfilment (Pettorelli et al., 2014b). However, the lack of consistent global measures of progress has been compounded by the diversity of methods in use and by limited cross-party cooperation (Skidmore et al., 2015). The concept of Essential Biodiversity Variables (EBV) was proposed as a means of coordinating global action, and to provide common criteria for monitoring change (GEO BON, 2018; Pereira et al., 2013). They include variables such as species distribution, abundance and community composition.

To ascertain whether policies can reduce, or even reverse, the current decline in biodiversity, it is necessary to be able to detect changes in the occurrence or abundance of many species at large scales (Ferrarini et al., 2021). An appreciation of the significance of community structure has prompted a shift in modelling approaches, with an increasing focus on the influence of environmental conditions on entire plant communities rather than individual species (Ferrarini et al., 2021). Joint species distribution models provide a generalised understanding of how entire communities vary in response to abiotic and biotic environmental factors. Furthermore, many models now incorporate a spatial component, given that the distance between points, despite similarities in the abiotic environment, may influence the structure of the communities in question (Leibold et al., 2022). Selecting relevant environmental variables and accounting for their natural variation when modelling species' distributions is an important step (Dormann, 2007; Mod et al., 2016) in detecting the impact of anthropogenic effects, or variations in range brought about by climate change (Xu et al., 2019). Furthermore, species distributions are not only influenced by the abiotic environment, but also by biotic interactions between them. By analysing entire communities, studies can gain a more complete picture (Wisz et al., 2013). However, the efficacy of species distribution models is heavily dependent on the quality of the count and environmental data available.

Field surveys are constrained by time, money, and site accessibility (Skidmore et al., 2015), which in turn can influence survey design, sample size, and the scale of collection (Rocchini et al., 2016). It was recognised that the increasing spatial and radiometric resolution of Earth observation (EO) satellites, in conjunction with the continuity of long-term missions such as Landsat, provide a key opportunity for improving ecological monitoring (Wulder et al., 2016). Variation in surface reflectance is attributed to variations in ground substrate and differences in the pigmentation, chemistry and structure of individual plants and the communities they create (Frye et al., 2021; McDowell et al., 2015). The utilisation of EO data is becoming increasingly prevalent (He et al., 2015), and platforms such as Google Earth Engine facilitate accessibility to and analysis of EO imagery (Gorelick et al., 2017). However, the amount and types of data available are growing at a much faster rate than their utilisation (Regos et al., 2022), creating a gap between remote sensing and environmental modelling. Given that EO platforms are effectively observing ecosystems directly, it can be expected that reflectance values will be associated with field-level observations of ecosystem change. The use of open-access global satellite imagery offers the potential to identify more subtle variations in the landscape than may be identified by standard categorical or coarse scale environmental variables. This could assist in improving the identification of transitions between vegetation communities. The processing of remotely sensed products should then permit the inference of EBVs, including species richness, distribution, abundance and community composition (Jetz et al., 2019; Pereira et al., 2013; Skidmore et al., 2021).

While few species are observable directly or individually by Earth observation satellites (Bush et al., 2017), the benefit of EO in predicting species distribution has been demonstrated in cases where a single species is particularly dominant, making it distinct spectrally (Xing et al., 2021). Unfortunately, in the majority of cases, pixel values of surface reflectance represent a mixture of many plants, which consequently results in the dilution of the chemical or physical changes resulting from individuals (Lausch et al., 2016). The signal-to-noise ratio may be further confounded by collective vegetation responses to seasonal phenology, as well as within-season responses to environmental conditions such as changes in soil moisture (Regos et al., 2022). Finally, while monitoring should encompass the full spectrum of biodiversity, a significant proportion of species are rare, meaning that they contribute little or no information to the spectral profile of a satellite pixel. Nevertheless, the combination of the extensive coverage offered by EO imagery, with the detailed observations achieved by *in situ* field surveys, could enhance our calculations and provide a calibrated view of biodiversity trends.

This study uses historic vegetation monitoring data from the Greater Cape Floristic Region (GCFR) of South Africa, one of the most florally diverse areas on the planet. Some studies have been conducted in this region which combined topographic, climatic and limited spectral data to classify the GCFR into zones of homogeneous vegetation groups or to measure land degradation by calculating proportional vegetation cover (Bell et al., 2023; Van der Merwe et al., 2008a, 2008b). It is our contention that surface reflectance at the point and time of field surveys will more closely represent the status of the environment and the plant community than do static environmental variables derived from digital elevation and long-term climatic models. It is therefore anticipated that the inclusion of spectral reflectance will enhance the explanatory power of ecological models of the GCFR, which are otherwise based solely on static topographic and climatic variables (Chauvier et al., 2021). Furthermore, given the long-term and repetitive nature of satellite imagery, we expect that its inclusion will support the detection spatial delineation of changes to plant communities over time. To assist in monitoring progress towards international goals and to inform conservation action, we first test the hypothesis that the inclusion of EO data improves predictions of species distribution and abundance, relative to models based solely on static environmental variables. We then use EO to improve on the static and broad categorical delineation of plant communities (Dayaram et al., 2019; Vlok et al., 2005), by modelling and mapping plant communities at the species level across a conservation area, examining how they vary spatially and identifying where compositional change occurs over time.

2.3 Materials and methods

2.3.1 Study Area

The Greater Cape Floristic Region is of significant ecological importance, and includes two of only 36 global biodiversity hotspots, the Fynbos and Succulent Karroo biomes (Conservation International, 2023; Myers et al., 2000). The combined flora of the region comprises over 9,000 species, some 6,200 of these being endemic. The region has been heavily fragmented and degraded by invasive species, land clearance, climate change, and stock farming (UNESCO, 2021), with over 1,700 plant species now being listed as threatened, and 3,000 of conservation concern (Slingsby et al., 2017; UNESCO, 2021). This project combined data from five studies that were undertaken across the area between 2000 and 2019 (Figure 2-1, Table 2-1). The studies were conducted primarily within the Succulent Karoo, with sections of Fynbos interspersed, and some minor overlap with Nama Karoo to the north-east. In this project, the term survey refers to a single count of a single plot or transect. The majority of studies only surveyed plots on a single occasion, but at Sanbona, a private nature reserve, between 20 and 48 plots have been surveyed annually between 2004 and 2019, resulting in a total of 614 surveys being conducted.

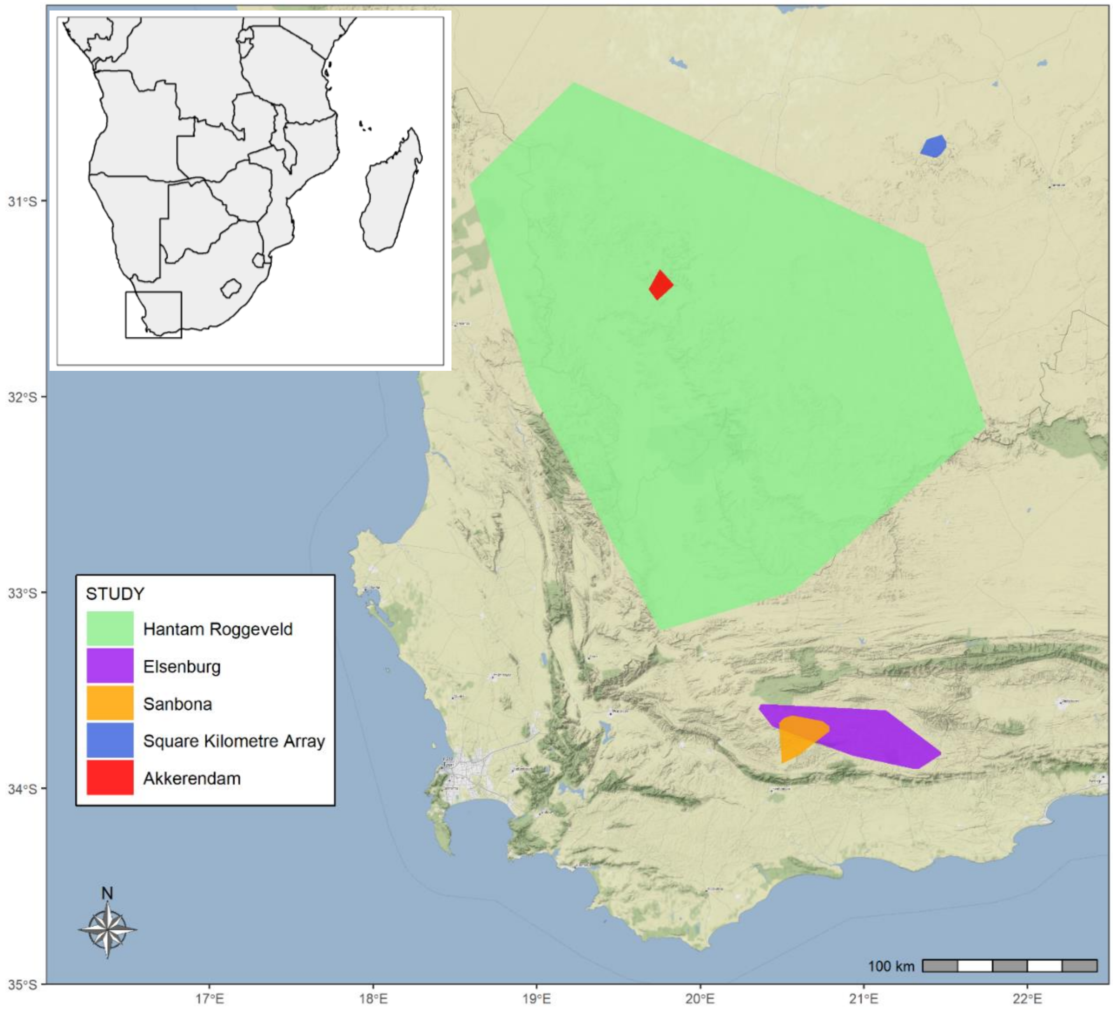

Figure 2-1: Location and range of the study areas in South Africa (details in Table 2-1).

Table 2-1: Details of vegetation studies used in this analysis.

Study Name	Year/s of Survey	Number of Plots/Surveys	Sampling Method	Plot Size	Number of Species Identified	Publication
Sanbona	2004-19 (repeated annually)	51/614	500 Point Line Intersection	500m²	208	Un-published
Elsenburg	2010-12 (counted once)	156/156	500 Point Line Intersection	500m transect	153	(Saayman et al., 2016)
Hantam-Roggeveld	2004	390/390	Braun-Blanquet	10-20m²	292	(Van der Merwe et al., 2008a, 2008b)
Akkerendam	2013	100/100	Braun-Blanquet	10-20m²	303	(van der Merwe and Hoffman, 2019)
Square Kilometre Array	2015	180/180	Braun-Blanquet	10-20m²	99	(van der Merwe, 2020)

2.3.2 Species data

In order to combine data from each study, and retain information about species relative abundance, the Braun-Blanquet categories (one, a few and cover of <5%, 5-25%, 25-50%, 50-75%, >75%), which were used in some studies, were converted to counts based on the median value of their class, and re-scaled to match the 500 point counts as used by other studies to combine data from different survey methods (Podani, 2006; van der Maarel, 2007). Only observations of species that were unambiguously identified were included in this study. Identified species were subsequently cross-referenced with the World Flora Online taxonomic reference list in order to consolidate all scientific names. This process was conducted using the R package "*WorldFlora*" (Kindt, 2020). The majority of the resulting 667 identified species were rare, and we elected to restrict analysis to species observed in ~1% of surveys of the dataset being analysed (15 surveys for the full set and 5 for the Sanbona subset), which falls within the range of the minimum number of recorded observations required for modelling as set out by J van Proosdij et al., (2016). Consequently, 211 species, were included in the regional analysis and 126 (of 208) in the Sanbona analysis.

2.3.3 Environmental data

In this study, we differentiate between standard "static" geographic data, which describe long-term mean climatic conditions and topographic characteristics, and "dynamic" EO imagery, which are continually updated. Five static predictor variables of each survey site were included in analysis; elevation, slope and aspect were obtained from the Shuttle Radar Topography Mission dataset (Farr et al., 2007), and mean annual rainfall and mean annual temperature were obtained from Worldclim (Hijmans et al., 2005).

There are numerous ways in which changes to vegetation composition and cover may alter their spectral reflectance, and the magnitude of these differences may vary both annually and seasonally (Blackburn and Milton, 1995; Inoue et al., 2008; Lausch et al., 2013). The selection of reflectance metrics used was not based on a specific *a priori* hypothesis regarding targeted taxa. Instead, the aim was to describe all axes of spectral variation within the study region. Few EO platforms have captured imagery consistently since 2004 and this study focused on six surface reflectance bands from blue to short-wave infrared that are available from Landsat 5, 7 and 8 (Wulder et al., 2019). Seven spectral indices which were considered to reflect important ecological and environmental changes in surface conditions were also calculated. Click or tap here to enter text.Click or tap here to enter text.Click or tap here to enter text.The data for the static and dynamic layers were downloaded from Google Earth Engine and their details, along with the derived indices, are presented in Table 2-2.

Landsat 7 imagery has known scan errors, where approximately 22% of each image lacks data. However, the land surface affected is not constant due to changes in the satellite's orbit in each pass, and errors are spread in strips a few pixels wide per image. The integration of multiple images can reduce the impact of these gaps, and the use of Landsat 5 or Landsat 8 images in conjunction with Landsat 7 for all but one year (2012) was considered sufficient and makes the use of the same process simple to follow in future studies (Claverie et al., 2018).

Table 2-2: Environmental and reflective variables used, including the vegetation indices calculated for use in analysis of Karoo plant data.

Category	Name	Description	
Climate	Mean Annual Temperature	Degrees C	
	Mean Annual Precipitation	Millimetres	
Topography	Elevation	Metres above mean sea level	
	Slope	Degrees from horizontal	
	Aspect (4 Categories)	45° either side of North, East, South, West	
Landsat	Blue	0.45 - 0.52 µm	
	Green	0.52 - 0.60 µm	
	Red	0.63 - 0.69 µm	
	Near Infra-Red (NIR)	0.77 - 0.90 µm	
	Short wave Infra-Red (SIR)	1.55 - 1.75 µm	
	Medium wave Infra-Red (MIR)	2.08 - 2.35 µm	
Indices	NDVI (Normalised Difference Vegetation Index)	(NIR-Red) / (NIR+Red)	(Rouse et al., 1973)
	SATVI (Soil Adjusted Total Vegetation Index)	(SIR-Red) / (SIR+Red) * (1.5) - (MIR/2)	(Marsett et al., 2006)
	MSAVI2 (Modified Soil Adjusted Vegetation Index 2)	((2NIR+1) - √((2NIR+1)^2 - 8*(NIR-Red)))*0.5	(Qi et al., 1994)
	MCARI2 (Modified Chlorophyll Absorption in Reflectance Index 2)	(1.5 * (2.5 * (NIR-Red) - 1.3 * (NIR-Green))) / √((2NIR+1)^2 - (6NIR-5 * √(Red)) - 0.5)	(Haboudane et al., 2004)
	NDWI (Normalised Difference Water Index)	(Green-NIR) / (Green+NIR)	(Mcfeeters, 1999)
	NDII5 (Normalised Difference Infra-Red Index 5)	(NIR-SIR) / (NIR+SIR)	(Hardisky, Klemas and Smart, 1983b)
	NDII7 (Normalised Difference Infra-Red Index 7)	(NIR-MIR) / (NIR+MIR)	(Chuvieco et al., 2002b)

To provide spectral metrics that are specific to the period of the vegetation surveys (hence dynamic), we summarised values from all EO images captured in the 12 months prior to a survey. For each survey point, the yearly maximum, minimum, mean and standard deviation, and the quarterly means, of all 13 reflectance values (six bands, seven indices) were calculated for pixels within 30 metres of the plot coordinate. A total of 104 spectral variables were thus calculated for each individual plot survey, comprising 52 annual variables (4 summary statistics x 13 spectral metrics) and 52 seasonal variables (1 statistic x 13 spectral metrics x 4 seasons). This high-dimensional description of the local land surface contains too many predictors to include in statistical analysis and will contain correlated redundant information. To retain information about the spectral differences among sites within fewer covariates, the EO variable stack was reduced to a smaller set of sparse canonical components using the *sgdm* package (Leitão et al., 2016), as detailed in Leitão, Schwieder and Senf (2017). This method is primarily used with zero-inflated data and differs from principal component analysis (PCA) by maximising the correlation between components and response variables, rather than the amount of variation explained in the data. Canonical components always condense data, with the number of components chosen varying between one and one fewer than the number of variables to be condensed. Fewer components retain less variation from the original data, unlike PCA, which always produces as many components as variables and retains 100% of the data's variation. For each dataset, preliminary analysis involved creating three, five and 30 canonical components, with WAIC values showing no significant differences between models fitted with these groups (14.1, 14.0 and 14.0 respectively).

In all datasets, collinearity was observed between some canonical components (Figure 2-2), though the colinear components differed between datasets,, and the dominant reflectance bands represented in each component varied. While collinearity between variables can reduce the interpretability of individual effects, it does not affect mean responses or predictions to new observations (Kutner, 2005). As the primary goal was to predict community composition rather than interpret the influence of individual canonical components, and to preserve maximum variation in the predictor data without overfitting the model, the observed collinearity was not considered to be detrimental. All subsequent analyses using EO data were conducted with models incorporating 10 variables: five static and five dynamic EO variables.

2.3.4 Statistical analysis

Vegetation composition was modelled using joint species distribution models (JSDM) in the R package *Hmsc* (Tikhonov et al., 2021), a particularly powerful example of the JSDM modelling framework (Norberg et al., 2019). Modelling rare species frequently yields unsatisfactory results due to the paucity of available data but *Hmsc* addresses this limitation by using spatial autocorrelation, and interspecific relatedness to improve inferences of rare species by their associations with more prevalent species. As the distribution of all species were highly zero-inflated, we adopt a hurdle-model approach, that fits a logistic model to presence-absence (P/A) data and subsequently models the changes to species abundances (log transformed) conditional on species presence (Ovaskainen and Abrego, 2020). The JSDMs were fitted with four chains of 10,000 iterations (the first 2,500 iterations were subsequently discarded as burn-in) and thinned to every 15th iteration so that the final posterior included 2,000 samples. The conditions under which dynamic EO information might offer the greatest benefit to static variables were identified by comparison of eight models that contrasted i) training data that were structured spatially (all data points) vs. temporally (Sanbona data points), ii) whether responses were fitted to presence-absence or abundance information, iii) whether models were fitted with only static covariates or also included dynamic EO variables. Spatial models were fitted using plots from a single survey of all sites. To avoid including the temporal structure of the Sanbona dataset within the spatial models, the average of 16 models, each using only one year of Sanbona data, were used to evaluate the spatial models.

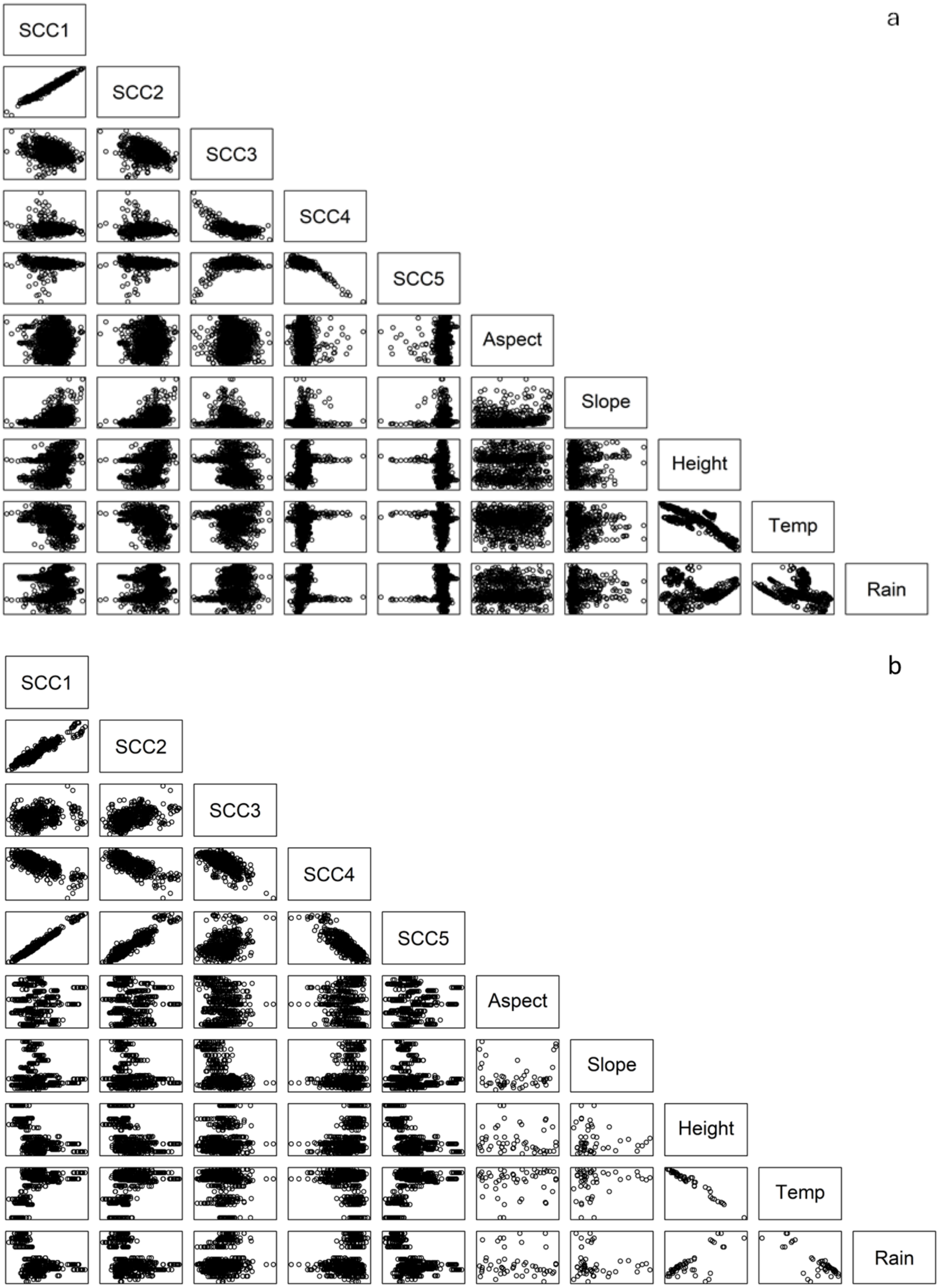

Figure 2-2: Pair plots of all predictor variables used in; **a** spatial data and **b** temporal data models. Canonical components were created independently for each set and the range of values of environmental data were different in both.

Explanatory power was evaluated by the mean R^2 across all species for abundance models, and the mean coefficient of discrimination (Tjurs R^2) for occurrence models. The proportion of variance attributable to individual explanatory variables in each model was identified using variance partitioning. Performance of the final temporal models was further quantified by conducting 6-fold cross-validation. Predictive power was tested by fitting new models to the Sanbona dataset, but excluding 2010/11 data, the same two years the adjacent Elsenburg study was conducted. The newly developed model was employed to predict the composition of sites in the Elsenburg study as well as the Sanbona sites for the two years that had been excluded. This enabled an assessment of the predictive capability of species composition at times and locations outside the training dataset. Only predictions for species that were shared between the training and test datasets could be assessed, which included 98 dominant species within Sanbona and 80 species in the Elsenburg sites. Prediction accuracy was evaluated per species using the coefficient of discrimination (Tjurs R^2), and for the entire community using Sorensen and Bray-Curtis dissimilarity indices for occurrence or abundance respectively.

To illustrate the potential for using EO data to scale up outputs from such models to inform policy, plant community composition was predicted across the Sanbona nature reserve to points centred 100 metres apart. Compositional differences between all points were described using a Sorensen similarity matrix, which was then translated into two ordination axes, which were further partitioned into red, green and blue channels and mapped using the 'recluster' package (Dapporto et al., 2020). In addition to spatial variation, composition may oscillate among related states or display directional change over time (Maliniemi et al., 2019). Community composition was calculated across the area for each year from 2004-2021, and year-to-year compositional differences for each point recorded. The potential for observing directional shifts in composition by identifying regions that were predicted to have experienced greater long-term (2004-2021) change than the standard deviation of year-to-year variation were illustrated.

2.4 Results

For the spatial data, incorporation of EO variables in species presence/absence models resulted in a marginal 1% improvement in the variance explained. Although the increase is small, EO variables accounted for 64% of the explained variance in the model, thus reducing the relative importance of the static and spatial covariates. In contrast, in species abundance models, EO variables increased the variance explained by 30% for the same spatial dataset, and represented 60% of the explained variance.

For the temporal dataset, models achieved far higher explanatory powers than for the spatial dataset when fitted with either presence/absence or abundance data. The addition of EO increased variance explained in presence/absence models by 3%, whereas the explanatory power of abundance models was increased by 7%. In temporal models, as in spatial models, the proportion of the variance explained by the EO covariates was greater than that added by them, at 17% and 34% for P/A and abundance, respectively. Figure 2-3 illustrates this phenomenon, showing that 83% of the variation in JSDMs fitted with temporal data of species presence/absence is driven by static covariates and spatial factors. In contrast, in the temporal abundance model, the dynamic EO covariates gain importance, primarily by absorbing variance that had previously been attributed to residual spatial covariance. The explained R^2 for each of the eight models as well as the proportion of the explained variance attributed to EO is given in Table 2-3. It is important to note that the Tjur's R^2 used for presence/absence (P/A) models is centred around zero, unlike the traditional R^2 as used for the abundance model, which is centred around 0.5. Consequently, a Tjur's R^2 value of 0.25 infers a reasonable ability to differentiate between presence and absence, and could be interpreted similarly to a traditional R^2 of ~0.45.

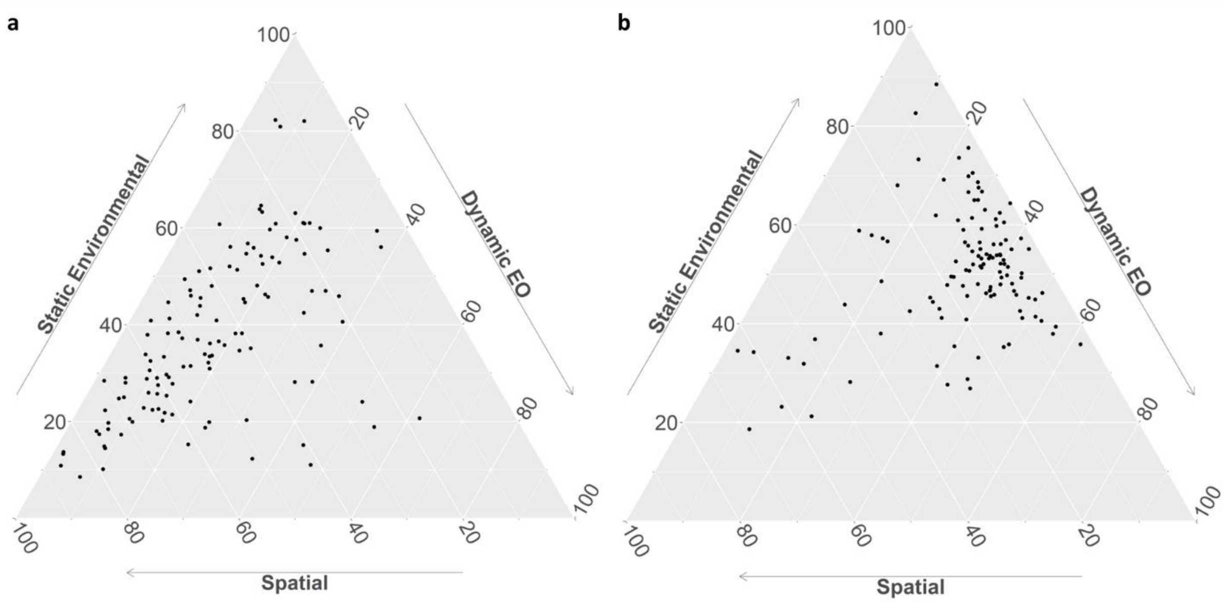

Figure 2-3: Proportion of variance explained for each species determined by spatial, static environmental and dynamic EO-based covariates when fitted with temporal data of **a**) presence/absence, and **b**) abundance information. The static environmental variables include elevation, aspect, slope, mean annual temperature and mean annual rainfall; and the dynamic EO variables are the five canonical components.

Table 2-3: The explanatory performance of models used in analysis, indicating the dataset and variables used, and whether the model was P/A (Presence/ Absence) or Abundance. The Static Variables column describes the explanatory power of models fitted with only static variables. The right-hand column indicates the proportion of the Total R^2 that is accredited to EO by variance partitioning, highlighting the disparity between the amount of explained variance they add and how much they represent.

Data Set	Data Type	R^2 using Static Variables	R^2 Added by EO data	Total R^2	Percent of R^2 accredited to EO data
Spatial	P/A	0.23	0.01	0.24	64%
	Abundance	0.23	0.30	0.53	60%
Temporal	P/A	0.45	0.03	0.48	17%
	Abundance	0.51	0.07	0.58	34%

When using Bray-Curtis dissimilarity values to quantify the accuracy of predicted against observed community structure, lower values are better, with zero representing a perfect prediction with no dissimilarity, and one representing no overlap in community composition. The mean dissimilarity in community composition between predictions made by the temporal presence/absence excluding 2010/2011 JSDM and the 80 Sanbona surveys conducted over 2010/11 was 0.54 (± 0.14 SD). In contrast, the mean dissimilarity between predicted and observed communities for the 156 sites at the neighbouring Elsenburg study was 0.72 (± 0.09 SD). However, when the same analysis was performed using abundance data, the mean dissimilarity with Sanbona decreased to 0.20 (± 0.06 SD) and to 0.35 (± 0.12 SD) at Elsenburg sites.

The predicted variation in communities across the Sanbona nature reserve is illustrated in Figure 2-4, and the vegetation types delineated on the national vegetation map (SANBI, 2019) exhibit some degree of correspondence. Our study suggests heterogeneity within the vegetation that encompasses both continuous, subtle changes in composition, and the presence of distinct vegetation types, which are repeated across the landscape. It was also predicted that plant communities underwent changes in compositional structure over time, as illustrated in Figure 2-5. The standard deviation of the year-to-year differences is illustrated, with darker areas representing greater variation in community structure during this period. This variation is significantly correlated with slope and elevation, with the lower, flatter areas experiencing greater variation ($R^2 = 0.17$, $p < 0.001$). In some areas, the shift in composition between the first and last years of predictions showed a difference exceeding the mean annual variation, suggesting a possible directional trend in compositional change. This trend was subsequently mapped in Figure 2-6.

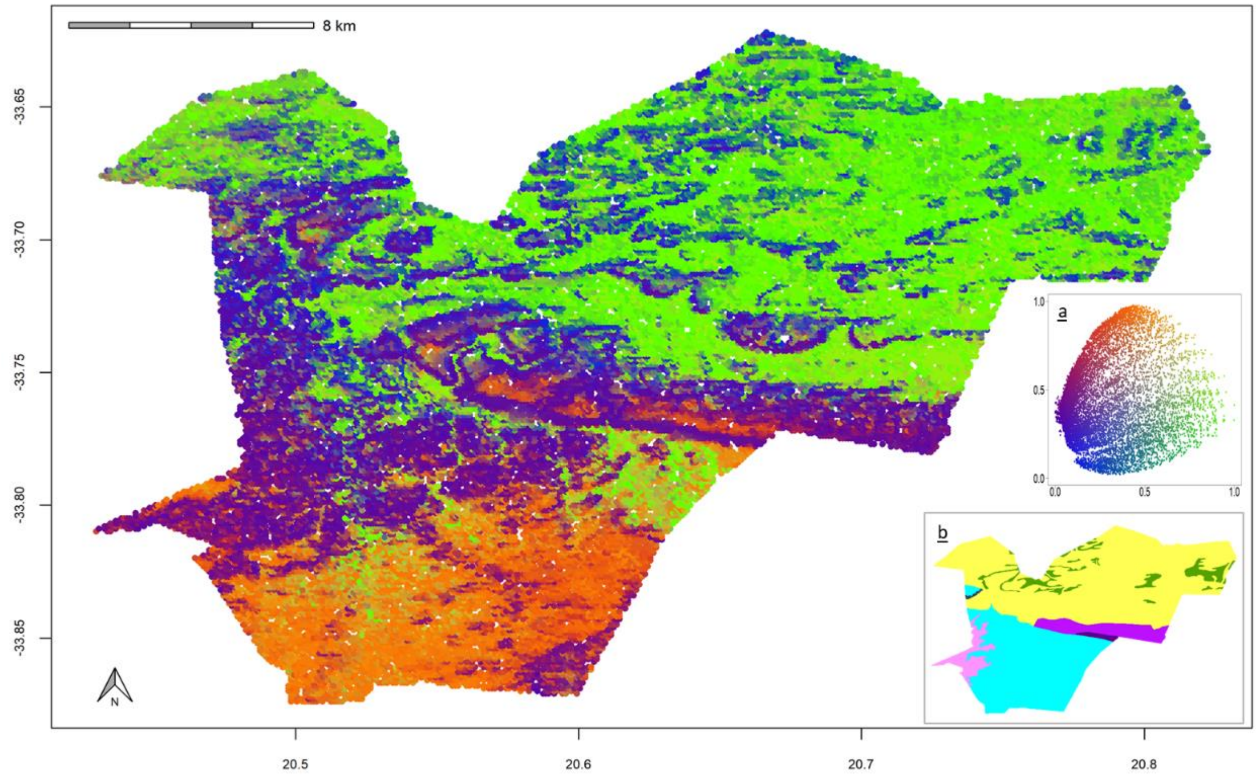

Figure 2-4: Variation in plant community composition across Sanbona. Similar colours indicate similarity in community composition. Insets a indicates community similarity by the represented colour. A difference of 1 indicates entirely different communities and a difference of 0 indicates identical communities. b shows the National Vegetation Map covering Sanbona and represents bioregions of communities with similar vegetation and abiotic features (SANBI, 2019).

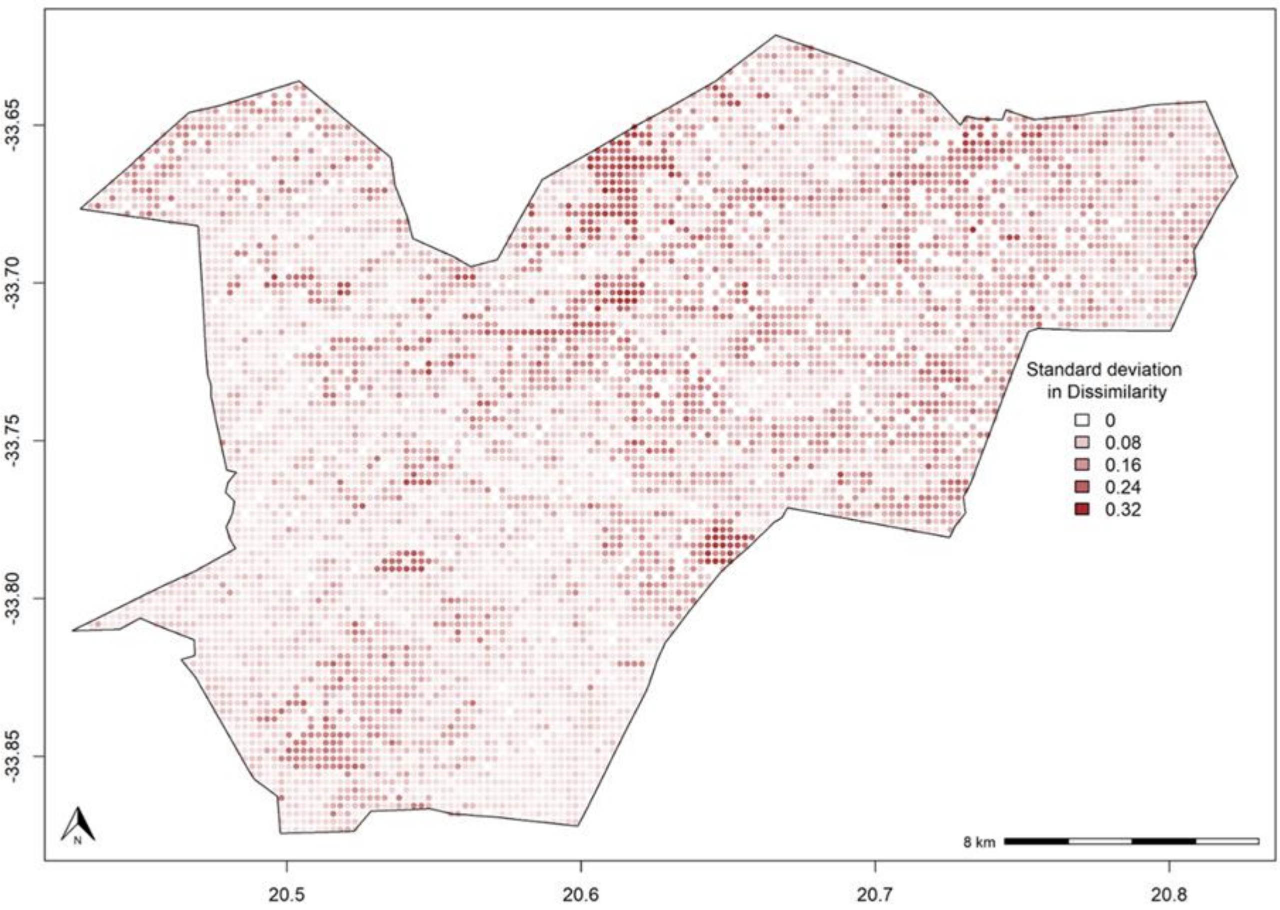

Figure 2-5: Standard deviation of dissimilarity per point over time of predicted vegetation communities in Sanbona. Dissimilarity is measured as the variation of an individual point between each two consecutive years and darker areas show where vegetation cover is predicted to have experienced the greatest year to year variation.

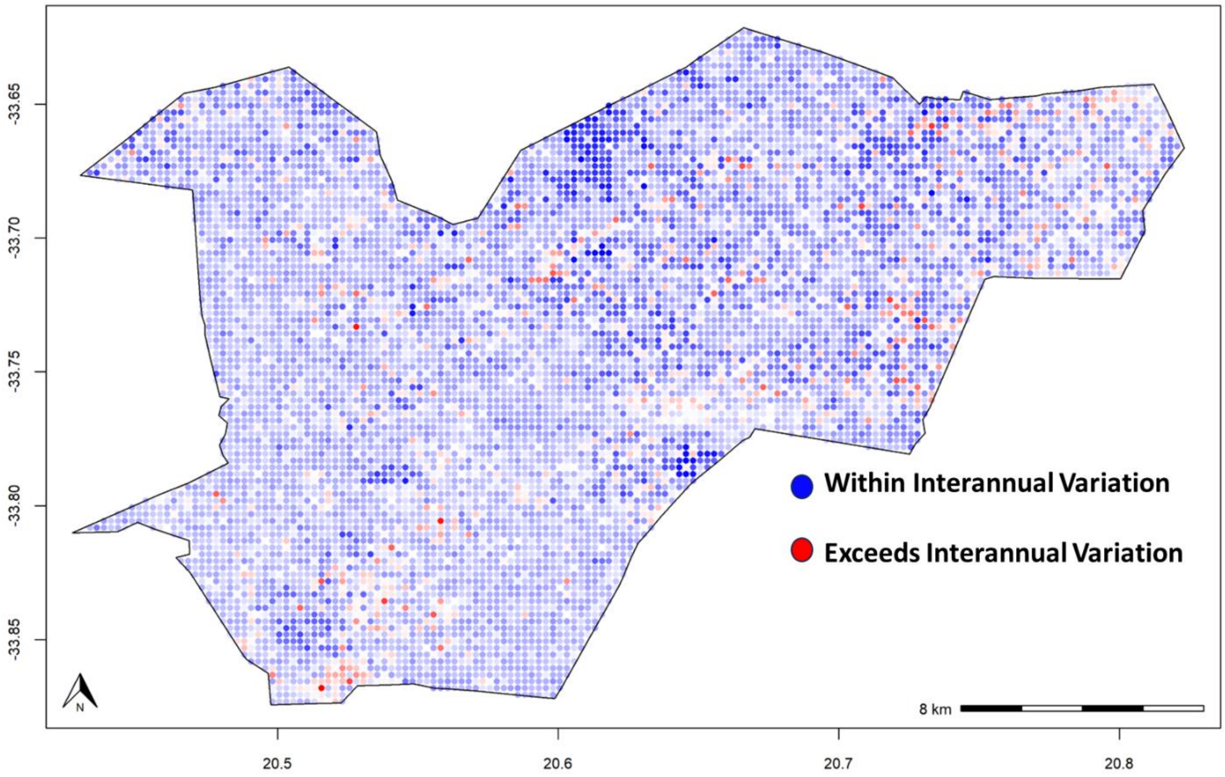

Figure 2-6: Dissimilarity in predicted vegetation communities in Sanbona per point over time. How the difference in communities between the first and last years of study vary from the standard deviation of dissimilarities per point across all 18 years 2004-2021. Blue points highlight where the difference between the first and last years is less than the standard deviation across the years, i.e. the measured difference is within a normal range. Red shows where the difference between the first and last years is greater than the standard deviation across the years, i.e. the difference measured is outside the normal range and may show where there is a directional change in community composition. Colour intensity infers the extent of difference between the first and last years and the standard deviation across years. Darker, more intense colours represent a greater difference between the two. Intense red represents areas where there is a greater difference and thus chance of community shift, and darker blue represents minimal change between the first and last years in an area that has a high range of change, inferring general stability within the community. Less intense colours indicate a low standard deviation, and/or a measured difference closely related to the standard deviation.

2.5 Discussion

Monitoring species distributions, abundance, richness and community composition is essential for achieving both local and global environmental targets. This study evaluated whether dynamic, satellite Earth observation data could enhance joint species distribution models, which relied on static topographical and climatic data to describe those ecological variables. Specifically, we aimed to determine whether satellite imagery could be used to improve on static vegetation map classifications by mapping plant communities at the species level, and where communities vary spatially and temporally across a conservation area.

Our findings show that the dynamic nature of EO data facilitated the observation of change over time in community composition, which would not have been possible by the use of static environmental data. However, EO variables did not substantially increase the explained variation in species distribution. However, EO data proved highly valuable for explaining variations in species abundance, particularly when assessing abundance changes across the entire spatial dataset. Furthermore, EO data accounted for a much larger proportion of the explained variance than they added. This improved predictability over large areas due to the measurable nature of EO reducing the contribution of the estimated latent variables used in spatial autocorrelation.

2.5.1 The influence of static environmental variables

Correlating species occurrence with topography and climate describes niches and the communities expected to be found in each (Pollock et al., 2014). Towards the edge of bioclimatic gradients, more complex topography and extremes of temperature and aridity can increase the number of available niches and influence the rate at which community composition changes (Ferrarini et al., 2021; Guerin et al., 2019). The Greater Cape Floristic Region is characterised by a climate that varies in terms of temperature extremes and particularly in the seasonality and amount of rainfall. It also features topographic diversity, encompassing plains, mountain ranges, and areas with sharply undulating peaks and valleys. In contrast to other studies in Mediterranean to semi-arid environments, we found that mean annual rainfall exerted only a moderate influence on species distribution (Caddy-Retalic et al., 2020; Guerin et al., 2019). Our results are consistent with those of Ferrarini et al. (2021), who found rainfall to be less influential than topography and temperature in classifying plant assemblages in the Swiss Alps. We found elevation and temperature to be the second most influential factors within our models. The inclusion of slope and aspect as proxies for topographic heterogeneity further increased the model's explanatory power, although they were the least significant predictors.

Current topographic-climatic niches do not consider historic drivers of community composition such as past climates, barriers to dispersal or competitive exclusion (Dan. L. Warren et al., 2014). This concept is potentially relevant in our area of study, as despite its topographic and climatic variation, Frye et al. (2021) found that plant communities changed relatively uniformly across the landscape. It has been suggested that the diversity of vegetation found in the Succulent Karoo may be a result of localised speciation driven by fine scale variations in soil and localised seed dispersal techniques, with the plant's sensitivity to environmental conditions and limited dispersal being further enhanced by their generally small size (Boucher et al., 2017; Ellis et al., 2006; Musker et al., 2021; Parolin, 2006). Our findings further support this, with spatial autocorrelation playing a greater role than topography and climate in explaining species distributions.

It is anticipated that the abiotic environment will exert a different influence on abundance than on occurrence (Mitchell et al., 2017). For a species to occur, it is presumed that local conditions allow for its survival, but once established its abundance is governed by how closely local conditions meet its ideal requirements. Our findings support this theory as abundance is most strongly associated with elevation and temperature rather than the distance between points found by occurrence. It is expected that temperature and elevation will be correlated (Runke et al., 2022), but the climatic data used in this study were recorded at a resolution of 930m^2, which is far coarser than the 30m^2 resolution of the elevation data. Consequently, elevation across the spatial data may more closely correlate with localised temperature and rainfall compared to the broader mean values supplied.

The higher explanatory power of models using temporal rather than spatial data when only fitted with static variables may be partly explained by the temporal data containing one-third fewer species than the spatial dataset. Furthermore, the temporal data were measured repeatedly, which, through repetition, may strengthen correlations between species and their environment (Ross et al., 2023), and increase the probability of encountering rarer species, thereby more fully describing the true community composition and improving inferences of interspecific correlations in a given physical environment (Pinto et al., 2022; Zhang et al., 2014). The benefit of multiple surveys was thus emphasised when only using fixed environmental variables.

2.5.2 The influence of adding EO data

Although EO added minimally to distribution models, their inclusion in models fitted to spatial data greatly reduced the influence of all static covariates and spatial autocorrelation, implying a high level of redundancy between the static and dynamic sets of covariates. The rainfall experienced in the year of a survey influences the water content and reflective properties of soil and plants. Dynamic EO captures this variation at a 30 m^2 resolution within the season of change, better representing the temporal and spatial nature of rainfall than long-term climate data (Zhang et al., 2011; Zhong et al., 2024). Similarly, while slope and aspect do not change over time, the north and south slopes of a hill, which have the same and unvarying static characteristics, may reflect differently within and between years, due to seasonal variations in sun intensity and rainfall patterns, and the influence that these have on soil and vegetation (Kumari et al., 2020). In terms of spatial autocorrelation, our models reduce the residual variance found after accounting for the influence of environmental variables by creating latent factors that vary with distance. These factors are likely correlated with unmeasured environmental variables, but unless new sites are very close to existing survey sites, these latent variables are unreliable contributors to predictions beyond the training data (Ovaskainen and Abrego, 2020). Thus, the addition of EO data reduces the residual variance in our model, which would otherwise be explained by spatial autocorrelation. Furthermore, as EO variables are quantifiable across the landscape, they enhance our ability to predict communities across a broader landscape (Chalmandrier et al., 2022; de la Fuente et al., 2021). In contrast, our study showed that incorporating EO data significantly enhanced the explanatory power of the spatial abundance model. Remote sensing studies have successfully described and mapped the distribution of large or dominant plant species, in part because their particular chemistry, phenology or flowers dominate the local signal (Nawrocki et al., 2020; Pinto-Ledezma and Cavender-Bares, 2021), more than do the occurrence of many rare individual species.

In both temporal data models, the proportion of variation attributed to EO data was again greater than they added, but the increase was less than observed in spatial models. Although the influence of all static variables decreased marginally and evenly, we found that, unlike in spatial data models, spatial autocorrelation remained unchanged. This suggests there is little redundancy between EO and spatial autocorrelation when describing species distribution and abundance in our temporal data. In addition, despite multiple surveys being conducted, and the smaller area over which the surveys were conducted, fewer species were identified in the temporal data set in comparison to the full data set. This aligns with the previously discussed concept of localised speciation (Boucher et al., 2017; Musker et al., 2021), which over large areas should give rise to more species than the smaller area covered by temporal surveys. Thus, over a smaller area, distance is likely to play a lesser role in describing community turnover than do the environmental variables that influence niches.

Furthermore, EO provided insight into community change over time, a feat unachievable using static environmental descriptors. The correlation between areas that had experienced the greatest inter-annual variation (Figure 2-5) with low lying and low sloped land could indicate an association with drainage flood plains, aligning with historical land use in the region, which frequently farmed crops or overgrazed those areas, leaving them with little vegetation cover (Van der Merwe et al., 2008b). These areas will vary in reflectance due to erratic seasonal rainfall and the proliferation of plants when rains come, more so than will areas covered with perennial vegetation (Bell et al., 2023).

2.5.3 Limitations and Opportunities

Surveys can struggle with scale relevance when determining species distribution and may not accurately represent true presence/absence (Gelfand, 2020). The Braun-Blanquet system used quadrats that were markedly smaller than Landsat pixels and the other survey methods. While this increases the likelihood that every species within a quadrat will be identified, the estimation of their cover area is less reliable for determining abundance. It is also likely that species present in the area but beyond the quadrat will be missed, thereby underestimating the true community composition. At Sanbona, survey plots were very similar in size to Landsat pixels. However, due to the method used, it is possible that smaller and less abundant species were missed (Perret et al., 2023), although the abundances of species were numerically recorded. Elsenburg surveys were comparable in count accuracy to those at Sanbona, though the linear structure of transects reduced their correlation with satellite pixels. While we considered each survey as a complete representation of the plant community at a site, given the species-rich ecosystem we acknowledge the limitations in fully capturing spatial and temporal patterns of biodiversity. Future studies will therefore benefit from surveying plots that are commensurate with satellite pixels and adopting a uniform count method relevant to the study at hand.

Surveys can also be subject to bias in site selection. The Elsenburg study sites were chosen to assess the impact of stock farms on vegetation, whereas the Sanbona sites were selected to represent a broad cross-section of landscapes, including but not limited to old, impacted farmlands. Although it seemed reasonable to compare community composition across studies given their proximity, the differing objectives behind each study may limit our ability to compare them. Furthermore, the SKA study was conducted in the Nama Karoo biome, which experiences different seasonal climatic patterns than the Succulent Karoo and Fynbos biomes. The smaller species list identified there results from the different climate, and results in a smaller overlap of species with other sites. Future studies would benefit from surveying sites that contain biomes or environmental impacts of interest, with sufficient overlap across the region of study to improve analysis and predictive interpolation.

The description of the GCFR was proposed as it encompasses two biomes (Fynbos and Succulent Karoo). The transitional nature of the area elicited debate about which biome certain areas should be classified (Born et al., 2007). Efforts have been made to delineate vegetation into groups and classes across the area, which have been mapped at various scales from national to regional. Despite the time and resources required to run computer analysis, our study shows that EO permits a more nuanced view of turnover amongst vegetation communities across the region. In doing so, a general mosaic structure of varying community compositions across the area becomes evident, and pockets of communities that differ from the dominant surrounding areas are highlighted. More importantly, EO can be used to identify areas that display extreme variation or directional change in community composition, and given training data that covers broad environmental gradients, can highlight the existence of eco-tones. These can serve as tools to guide conservation efforts, from local land management to regional and national initiatives, such as those outlined by (CapeNature, 2025; CEPF, 2009; SANBI, 2025).

2.5.4 Conclusion

The integrity of an ecosystem is derived partly from the interactions among communities at the landscape scale (Le Provost et al., 2022). Planning and assessing the conservation of hundreds of plant species requires a framework that can be efficiently scaled to inform priority locations and actionable interventions (Ferrarini et al., 2021). By integrating environmental and EO data with field survey data in joint species distribution models, it is possible to expand predictions of species richness, individual probabilities of occurrence, and community composition across a landscape, thereby enabling the prioritisation of resources for conservation. The results of this study demonstrate that EO greatly enhances our ability to explain the variation in species abundances at a landscape scale. Furthermore, EO can be used to predict and map communities across a large area and provide an understanding of how and where communities change, both spatially and temporally, which standard vegetation class maps do not reveal (Regos et al., 2022). With new satellite images available almost daily, this source of information could support near real-time mapping of vegetation communities across hyper-diverse, semi-arid environments, such as the GCFR, as well as other biomes where sufficient field data are available for calibration.

Code and data used in the analysis of this chapter are available at
https://github.com/AndrewCSlater/HMSC_analysis

3. Predicting ecological variation in a global biodiversity hotspot: linking metabarcoding to Earth observation

3.1 Abstract

To halt the decline in biodiversity and ameliorate climate change, it is vital that remaining tropical forests are conserved or managed sustainably, yet half of the world's forest reserves are experiencing a loss of functional and taxonomic diversity. Earth observation (EO) from satellite remote sensing provides free, consistent, and repeatable data at global scales, which can be used to infer variation in forest canopy composition across space and time. However, field data are required to understand how that variation relates to biodiversity status and trends. This study combined field surveys of insect biodiversity, processed using DNA metabarcoding, to test whether continuous data from EO could act as an effective indicator of ecological condition.

Gola Rainforest National Park (GRNP) in Sierra Leone is part of a global biodiversity hotspot and is subject to a Reducing Emissions from Deforestation and Degradation (REDD+) programme. The park is surrounded by a 4 km-wide, community-owned buffer zone in which REDD+ payments are intended to improve local livelihoods on the condition that the landscape is sustainably managed and biodiversity value maintained. Robust monitoring of ecological integrity and verifying compliance through field surveys is not considered viable across this area, and we therefore sought to investigate the potential for remotely sensed data sources to indicate effective buffer zone management.

Modelling of the 284 most frequently detected insect taxa demonstrated that the satellite-derived covariates explained community variation as effectively as field-measured vegetation data. When predicting to new locations, 55 taxa achieved area under the curve (AUC) values greater than 0.7 using EO data, while 49 taxa reached this threshold using field data. However, the taxa predicted well by each method only partially overlapped. Field data highlighted taxa responded to particular vegetation types, whereas EO data identified taxa associated with structural gradients linked to fragmentation.

The findings of this pilot study underscore the importance of high-quality ecological data obtained through field surveys, not only to assess biodiversity directly, but to train and validate models that link biodiversity to spectral reflectance. These models allow biodiversity to be predicted across larger areas, but the ecological meaning of these predictions in terms of habitat condition and management practices can only be reliably interpreted through ongoing field surveys that capture key structural and ecological attributes across the landscape. The rapid and efficient capture of such complex ecosystems is facilitated by the use of new technologies. The models developed here may serve to guide further field surveys and, in the future, support the strategic delivery of payments for sustainable management.

3.2 Introduction

Tropical forests are home to more than half of the world's biodiversity, making them some of the most biologically diverse areas on Earth (Bradshaw et al., 2009; Gibson et al., 2011), and more than a billion people depend directly on them for food, materials, and services (Lewis et al., 2015). Furthermore, tropical forests account for approximately half of the global terrestrial carbon uptake and up to two-thirds of the carbon sink of all forests (Hubau et al., 2020; Pan et al., 2011). Despite the expanding area of tropical forest falling under protection, half of the world's tropical forest reserves continue to experience a loss of taxonomic and functional biodiversity (Jenkins and Joppa, 2009; Laurance et al., 2012). The anthropogenic pressures surrounding forest reserves are often reflected within them (Laurance et al., 2012). Logging and agriculture are the most significant drivers of forest habitat change and loss. The practice of clear-cutting and commercial agriculture has the potential to completely destroy forest areas. However, in tropical Africa, selective logging is more common than clear-cutting and can lead to changes in the density and composition of understorey vegetation, which in turn can facilitate the cultivation of understorey crops such as cocoa (Potapov et al., 2017). These shifts in forest structure, whether linked to logging, subsistence farming or commercial crops, contribute to forest degradation, further species loss and significantly expand the area of human impact (Gibson et al., 2011; G. D. Lennox et al., 2018). Efforts to prevent further destruction and degradation of tropical forests must be adaptive and guided by up-to-date information. In order to learn from the successes and failures of the strategies employed, it is necessary to continuously measure environmental conditions across space and repeatedly over time (Pettorelli et al., 2014b).

The implementation of field surveys of vegetation is constrained by financial, spatial, and temporal limitations. Consequently, the use of EO is considered an important approach to improving vegetation monitoring, if not also for associated biodiversity (Skidmore et al., 2015). The calculation of categorical landcover and deforested areas from EO imagery enables temporal monitoring (Azzari and Lobell, 2017), but requires *a priori* assumptions about the importance of the defined classes and ignores the continuous nature of land cover. The continuous measurement of satellite-derived elevation and climate data has been used to infer the capacity for an area to grow forests (Fiandino et al., 2020), and to broadly describe elements of environmental condition or niches (Guisan and Zimmermann, 2000). More recently, LiDAR has proven to be a powerful tool for describing the three-dimensional structure of forests. However, its application at landscape scales remains limited, primarily due to high costs and the impracticality of multi-temporal surveys (Asner, 2007; Dubayah et al., 2020). An alternative approach is to apply the continuous variation of processed satellite pixel values over large areas, which can also be projected back in time using Landsat. The heterogeneity within a landscape, such as variations in soil, topography and disturbance, and the resulting vegetation communities found, combine to cause variation in spectral reflectance (S. D. Warren et al., 2014). The spectral variation hypothesis suggests that heterogeneity observed within images is related to the number of ecological niches present on the ground (Purdon et al., 2022), and that surface reflectance values can be used in models to predict community and structural variations (Csillik et al., 2019).

The number of available niches is positively correlated with species and functional biodiversity (Stein et al., 2014). The measurement of structural elements such as patch size, or canopy height, can provide useful indicators of forest condition and habitat fragmentation (Clark et al., 2021). However, while these measures can reflect aspects of ecological integrity, they do not directly describe diversity or community composition. A variety of taxa, including birds and insects, have been studied to determine how forest structure and disturbance influence their communities (Bregman et al., 2015; Parikh et al., 2021). Biodiversity responses to habitat changes are widely used as indicators of ecosystem health, and by monitoring changes in ecological communities, it is possible to assess whether conservation and management efforts are maintaining habitat quality (Niemi and McDonald, 2004). The richness and composition of insect communities have been demonstrated to respond to and reflect forest age, structure and composition, and plant diversity (Basset et al., 2012; Schowalter, 2017; Traylor et al., 2022; Zhang et al., 2016). Insect communities are vast and sensitive to habitat change, making them valuable biological proxies for assessing environmental condition. Advances in high-throughput molecular tools such as DNA metabarcoding now allow efficient identification of entire insect assemblages, facilitating their use in large-scale monitoring (Buchner et al., 2023). By linking insect diversity patterns to forest structural variation, it becomes possible to infer biodiversity responses across broad areas. Remote sensing-derived proxies of forest structure therefore offer a potential approach to monitoring habitat condition at landscape scales, but there remains limited understanding of how effectively these proxies reflect biodiversity itself. Addressing this gap is essential for developing effective monitoring and conservation strategies.

Our study focused on the 4 km wide REDD+ leakage belt, or buffer zone, surrounding the Gola Rainforest National Park (GRNP) in Sierra Leone. The REDD+ framework requires the development of sustainable livelihoods for the 122 communities situated within the leakage belt, in conjunction with the implementation of measures to limit deforestation (to avoid leakage and loss of stored carbon), and the maintenance of biodiversity value. Combining these disparate objectives within a single landscape requires the integration of outputs from carbon, biomass, land cover class and biodiversity studies. Such a system relies on the development of dependable and empirically founded methods for monitoring across landscape scales. By improving biodiversity monitoring at scale, this approach could ultimately support conservation planning and guide sustainable land-use decisions, including agricultural intensification strategies that minimise biodiversity loss.

This study uses aerial insect communities to describe biodiversity and explores how well aspects of forest structure, measured through both field surveys and EO proxies explain variation in these communities across the buffer zone of the GRNP. Specifically, we ask:
Does forest structure, as determined by field surveys, explain and predict spatial variation in insect communities across the buffer zone?

To what extent can proxies of forest structure derived from satellite remote sensing predict biodiversity patterns of insect communities across the same area?

To address these questions, we tested the hypotheses that variation in forest structure, as determined by field surveys, is associated with differences in insect communities, and that EO derived proxies of forest structure provide stronger predictive power for insect community variation than direct field measurements. The outputs of this analysis are used to create maps of insect community similarity and richness.

3.3 Materials and Methods

3.3.1 Study Area

The GRNP forms a large part of the remaining Guinean Forests of West Africa, a global biodiversity hotspot (Conservation International, 2023; Myers et al., 2000). Our study was conducted across a 40km stretch of the buffer zone surrounding the central block of the GRNP. The buffer zone is 4km wide and comprises a mixture of primary and secondary forest and areas that have been clear cut. Both forest types may experience selective logging, and a patchwork of land uses (patch type), including agroforestry cocoa, low-intensity oil palm and subsistence agriculture may be found in the forest under canopy as well as in the clearcut areas. A diverse array of beetles and low-flying insect species were collected using Malaise traps (Uhler et al., 2022) at 105 sites within the buffer zone, and 11 sites from within the GRNP Figure 3-1, during the months of November and December 2021. Trap sites were selected *in situ* to cover a broad range of forest and land use, and within a general survey area each sites were separated by at least 150m. A link to the raw data is provided at the end of this chapter. Each trap was left for a period of five days, and samples were collected and stored in 100% ethanol. To describe the environment within a 30m radius of each trap, the maximum canopy height, percentage cover of bare ground, grass, cocoa, lianas, and forest were estimated. Additionally, evidence of burning, the presence of oil palms, or tree stumps greater than 10cm or 30cm were recorded.

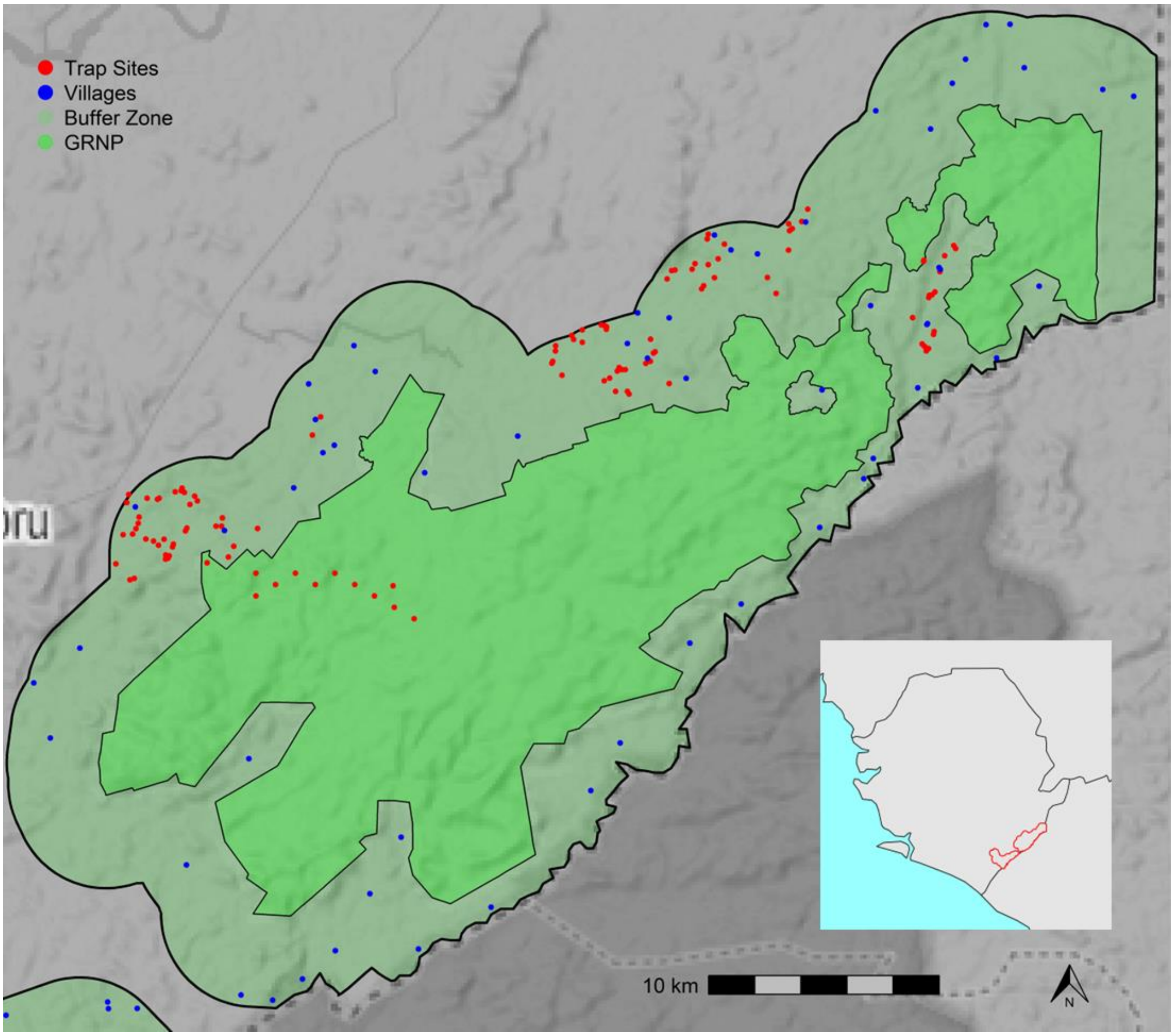

Figure 3-1: Area of study: The Northern and Central blocks of the GRNP and their buffer zones, with the location of traps and villages indicated. Inset: Location of the buffer zones around the combined northern and central blocks and the southern block of the GRNP within Sierra Leone.

3.3.2 Sample Processing

Malaise trap samples were processed using DNA metabarcoding (Piper et al., 2019) in order to identify operational taxonomic units (OTUs) based on similar DNA clusters identified. Each OTU was then compared to the Barcode of Life database to assign it to the lowest taxonomic level possible. A description of the sequencing process can be found in the appendices. Of the 3,869 OTUs recorded, 1,268 were identified to a taxonomic kingdom, with 1,152 of those being identified as insects. Our study analysed the 284 insect OTUs that were found in at least three sites, which was the suggested minimum for these models (Pichler and Hartig, 2021).

3.3.3 Remotely sensed data

A wide range of Earth Observation (EO) platforms have been employed to approximate forest properties in the past. As previously described, spectral properties are expected to respond to variations in vegetation types. To compare satellite arrays against each other, a single cloud-free surface-reflectance image, taken within one month of the field surveys, was obtained for Landsat-8 and Sentinel-2, and a cloud-free composite image for March 2022 from Planet. For each image, and dependant on the available reflectance bands, up to nine vegetation indices were calculated that highlight ecological properties of the reflected surface. The normalised difference vegetation index (NDVI), although prone to saturation over dense forests (Huete et al., 1997), is still able to determine between other vegetation gradients (Irteza et al., 2021) and is included as deforested areas are included and as a standard index for vegetation productivity. The perpendicular vegetation index (PVI) and modified soil adjusted vegetation index-2 (MSAVI2) highlight early regrowth from bare soil; the enhanced vegetation index (EVI), normalised difference red edge (NDRE), and green normalised difference vegetation index (GNDVI) highlight changes in dense canopy cover; normalised difference infrared indices five and seven (NDII5 and NDII7) react to variation in leaf moisture; and finally, the normalised burn ratio (NBR) and visible atmospherically resistant index (VARI) measure disturbance caused by recent burning and clear cutting, which also coincide with the field survey descriptors of evidence of recent burning or logging. It is recognised that frequent burning and the time since the last burn are also potential influencers of forest structure (Burivalova et al., 2015; Rappaport et al., 2022, 2018), but these variables were beyond the scope of this study and not included in analysis.

In addition to the spectral properties of individual pixels, spatial patterns in pixel values are also expected to convey information about forest structure (Rocchini et al., 2004). Grey Level Co-occurrence Matrices (GLCM) describe the pattern of co-occurring neighbouring pixel values of a single raster layer (Zhou et al., 2017). These matrices were calculated for each spectral band and vegetation index. GLCMs describe the pattern of co-occurring neighbouring pixel values of a single raster layer (Figure 3-2) within a moving window. A window size of approximately 90m was used for all satellite platforms (Landsat:3 x 30m pixels, Sentinel:9 x 10m pixels, Planet:17 x 5m pixels), which from visual inspection of satellite images, appeared to be an appropriate size for effectively covering a uniform patch type with minimal influence by neighbouring patch types on the window (Hall-Beyer, 2017). To describe the variation within GLCMs, the contrast, entropy, and mean values were calculated (Hall-Beyer, 2017; Haralick et al., 1973). GLCM calculations were performed in R using the “glcm” package (Zvoleff, 2020). The values for each spectral band, vegetation index, and GLCM layer were summarised as the mean pixel value and standard deviation of pixel values within a 30m radius of each trap location. The number of reflectance bands, and subsequently, the number of indices differ between satellite platforms. Consequently, the number of EO descriptors for each platform varies. Landsat-8, Sentinel-2, and Planet had 120, 150 and 80 EO descriptors respectively. A summary of the metrics for each satellite is shown in Table 3-1.

Figure 3-2: Examples of three theoretical forest patch types. All three have the same number of dark (valued 3)and light (value 1) green pixels, thus mean statistics would describe each block of 16 as medium green (value = 2) with equal standard deviation to colour. However, GLCMs value each pixel not by its colour but by the patterns of neighbouring values and would provide different values for each pixel and thus patch.

In addition to optical EO platforms, imagery captured by synthetic-aperture radar satellites such as Sentinel-1 is also used. These satellites emit polarised energy, which is influenced by the geometric structure and water content of the area being observed. The strength and polarity of the returned signal are measured, thereby providing insights into the surface structure. Studies have demonstrated the utility of radar imagery in the measurement of biomass and the differentiation of different forest types and areas of similar colour reflectance that have differing structural properties (Numbisi et al., 2019; Solberg et al., 2014). An ortho-corrected and processed image from Sentinel-1, , captured in January 2022, containing single and dual-polarised (Vertical-Vertical and Vertical-Horizontal) bands, was used. Similarly to the optical platforms, each band was represented by the mean and standard deviation of pixel values around the trap.

Finally, a contemporary snapshot may mask longer-term impacts, particularly in the GRNP buffer zone, which contains fragments of both original and recovering forest. We therefore also included the time since deforestation, calculated from historic Landsat data in the Global Forest Change Database (Hansen et al., 2013). Likewise, EO variables may fail to identify more subtle drivers of disturbance that relate to accessibility. Consequently, the distance to the GRNP boundary for each trap point was also included, with traps outside the GRNP being given a positive distance, and those inside, a negative distance.

Table 3-1: The colour bands available and indices created for each satellite platform.

Bands	**Landsat 8**	**Sentinel 2**	**Planet**
Blue	x	x	x
Green	x	x	x
Red	x	x	x
Red-Edge 1		x	
Red-Edge 2		x	
Red-Edge 3		x	
Near Infra-Red	x	x	x
Short-wave Infra-Red	x	x	
Medium-wave Infra-Red	x	x	
INDICES			
EVI	x	x	x
GNDVI	x	x	x
MSAVI2	x	x	x
NBR	x	x	
NDII5	x	x	
NDII7	x	x	
NDRE		x	
NDVI	x	x	x
PVI	x	x	x
VARI	x	x	x

Insect distribution is influenced by the climate of an area, and seasonal changes may influence the communities found, but variations in community composition between forest vegetation types have been shown to remain constant throughout the year (Zhang et al., 2016). While global climate change can also influence insect communities, these changes can be difficult to generalise and gauge from a single survey (Pureswaran et al., 2018; Subedi et al., 2023). Insect distribution has also been shown to have some correlation with elevation, although studies vary in their findings and it is thought that much of the correlation found can be ascribed to variations in sampling timescale, human disturbance and climates found at different elevations (Hodkinson, 2005; McCoy, 1990; Wolda, 1987). As we were specifically testing the ability of satellite observation to predict the distribution of insect communities, we felt the inclusion of climate and elevation could have detracted from this analysis, and they were excluded from our models.

3.3.4 Statistical Analysis

Insect communities were modelled using joint species distribution models (JSDM). JSDMs model all species concurrently, considering the effects of environmental gradients on species distributions, as well as spatial autocorrelation and correlations in pairwise-species cooccurrence (Leibold et al., 2022). To accommodate the large matrices generated by DNA metabarcoding, which are characterised by a high proportion of absences, we used the sparse-JSDM package "*sjSDM*" (Pichler and Hartig, 2021), utilising the recommended elastic-net and regularisation parameters (M. Pichler, personal communication, 20th June 2023).
To investigate the impact of forest habitats on insect communities, an sjSDM was fitted to the entire dataset using the 10 habitat descriptors that were measured *in situ.* Subsequently, the data were randomly divided into five groups and five sJSDMs were created, each withholding one of the five groups and being fitted to the remaining four. Each model was subsequently used to predict to the hold out group. AUCs were then calculated to enable a comparison of predictive performance with EO models.

The wide variety of methods used to process satellite imagery are indicative of a range of ways in which EO data may recover elements of forest composition or structure (Lausch et al., 2016). EO variables were designated one of four groups: raw reflectance bands, vegetation indices, GLCM of bands (GLCM-B) and GLCM of indices (GLCM-I). To address the high dimensionality of the EO data sets, which ranged from eight (8) to 155 per group, dimension reduction was applied to prevent overfitting in subsequent analysis. Sparse canonical components (SCC) were used to maximise the correlations between EO and ecological (survey) data, while condensing the EO data to a fixed number of variables. This approach ensured consistent dimensionality across models, allowing for direct comparison of their performances. The smallest group (raw spectral bands of the Planet satellite array) contained eight variables, which limited the number of SCCs to seven to ensure some degree of data compression while preserving statistical stability. Consequently, in preliminary analysis, all combinations of EO reflectance variables were reduced to seven SCCs to maintain uniformity in the number of predictor variables used in each model. The four Sentinel-1 radar values were independently reduced to one SCC. All SCCs were created using the *sgdm* package (Leitão, Schwieder and Senf, 2016). To exhaustively test the performance of EO variables, for each optical satellite platform, 120 JSDMs were fitted, covering all possible combinations of EO groups, combined with all possible combinations of the additional variables of radar, year of deforestation and distance to the GRNP (

Table 3-2). Model performances were measured by McFadden's pseudo-R^2, an option built into the sjSDM package, which was used to ascertain which satellite array and group of variables would be used for ongoing analysis. Subsequently, to find the most appropriate number of SCCs to use, models were fitted with between one and 20 SCCs, created using data from the best performing variable group and satellite platform. It was found that 10 SCCs most effectively fit to our survey data, and this number was used in all further analysis.

To test the sensitivity of model performance to the scale of measurement, ten optical SCCs were then created by processing reflectance values measured at different radii around each trap (30, 60, 90, 120, 150, 250 and 500m). In addition, one radar SCC was generated for each radius. The data sets for each radius were combined with the fixed values of the year of deforestation and distance to the GRNP, and then modelled independently. For each radius, the complete data set was randomly divided into five groups, and five models were fitted. Each model was fit to four groups, which was then used to predict to the fifth held out group. The AUC for each model was then calculated on a per species basis. The mean predictive AUC across all five models for each radius was then calculated, and the model with the highest predictive performance selected. The distribution of variables used in the final model are plotted in Figure 3-3. Correlation was found between the first two canonical components, but this was considered acceptable given that the goal was prediction rather than causal inference (Dormann et al., 2013; Doser et al., 2023). To minimise the potential impact of group selection, the AUCs for the final model were calculated four times, with a different random split of the data set used each time. The OTUs for which predictive performance was considered good (AUC>0.7) (Mandrekar, 2010), hereafter referred to as high-AUC OTUs, were identified and were subsequently used in further predictive analysis.

Finally, in order to better understand how much the EO variables related to the structure of the forest canopy, the same variables from the most effective EO model were employed to predict the ~30m point-level estimates of maximum canopy height, percent canopy cover, and foliage height diversity derived from the GEDI lidar platform (Dubayah et al., 2022). A random-forest model was constructed using the "*randomForest*" package with default settings (Liaw and Wiener, 2002) to analyse 2,448 GEDI points within the study area. The models were evaluated by the percentage of variance explained and the root mean square of residuals on out-of-bag samples. The most influential variables were noted. Simple linear regressions were employed to investigate the relationship between the probability of occurrence of each of the high-AUC OTUs and the maximum canopy height, percent canopy cover, and foliage height diversity estimated by GEDI. The direction of the significant relationships between OTUs and each of the descriptors was noted.

Table 3-2: The variables and groups used as covariates in models. The 15 combinations of reflectance variables were multiplied with the 8 possible combinations of other variables, giving 120 combinations to be modelled per satellite platform.

Reflectance Variables	**Other Variables**
Bands	No Additional Variable Used
Indices	Radar
GLCM from Bands	Deforestation Year
GLCM from Indices	Distance to GRNP
Bands + Indices	Radar + Deforestation Year
Bands + GLCM_B	Radar + Distance to GRNP
Bands + GLCM_I	Deforestation Year + Distance to GRNP
Indices + GLCM_B	Radar + Deforestation Year + Distance to GRNP
Indices + GLCM_I	
GLCM_B + GLCM_I	
Bands + Indices + GLCM_B	
Bands + Indices + GLCM_I	
Bands + GLCM_B + GLCM_I	
Indices + GLCM_B + GLCM_I	
Bands + Indices + GLCM_B + GLCM_I	

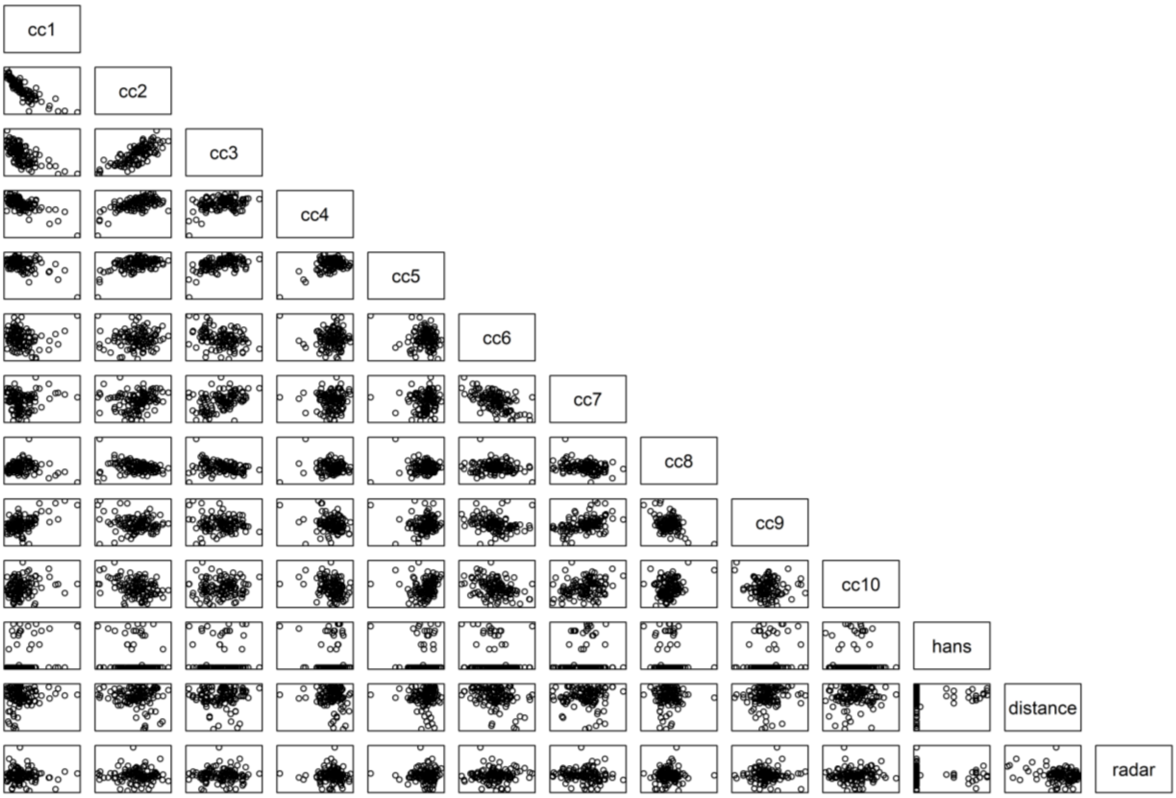

Figure 3-3: The distribution and relationship of the predictor variables used in the final model. These comprise ten Landsat canonical components (SCC1-10), Hansen year of deforestation (hans), distance to the boundary of the GRNP (distance) and the single Radar SCC (radar).

3.3.5 Prediction

The final EO model, based on the most effective satellite platform, variable group, buffer value and number of SCCs, was fitted to the full data set to predict the whole community across the central GRNP and its buffer zone. This was then used to map the distribution of the high-AUC OTUs. When predicting to larger areas outside the range of training data, environmental variables were clamped to lie within the range of the training data, with higher and lower values converted to training data’s maximum and minimum values (Anderson and Raza, 2010; Li et al., 2024).

To illustrate the variation in communities across the landscape, a community similarity index was calculated based on the probability of occurrence of all high-AUC OTUs per point. This was achieved using the “*Rtsne*” package (Krijthe, 2015), which produced two ordination axes. These axes were used to create colours for the mapping of community similarity across the area using the “*recluster*” package (Dapporto et al., 2020).

To use insect communities to infer biodiversity levels, species richness was calculated as the sum of all probabilities of occurrence. Additionally, to explore whether richness varied with forest structure, for each of the three measured structures, the richness of OTUs that exhibited significantly positive and negative relationships was calculated. Summing OTU richness in this way, may provide a guideline of where maximum richness exists at both sides of the spectrum in terms of forest structure, as described by the proportion of canopy cover the maximum canopy height, and the diversity of canopy height. These three metrics are indicators of both horizontal and vertical canopy complexity, which would be expected to vary along with changes in forest disturbance and succession (Iheaturu et al., 2024). The residual richness after subtracting the positive and negative relationship richness from each other may provide a useful tool in identifying areas for protection as they are likely to be high quality forest, or heavily disturbed forest where agricultural intensification will have the lowest impact.

All analyses were conducted using R version 4.1.0 (R Core Team, 2022) in R Studio (R Studio Team 2021)(R Core Team, 2022) in R Studio (R Studio Team 2021), and a link to the code and data used in our analyses is available at the end of this chapter.

3.4 Results

A total of 1,152 insect OTUs were identified in the 116 trap sites. Of these, 714 were only recorded once. Further analysis was conducted on the 284 OTUs that were recorded in three or more sites (Pichler and Hartig, 2021). Given that the mean OTU richness at each site was just 26.7 (SD=16.5), it is evident that the mean pairwise dissimilarity between sites was extremely high (Sorensen=0.99, Simpson=0.98). The near identical dissimilarity values indicate that observed differences in community composition are likely the result of complete turnover of OTUs, rather than one community being a nested, less rich subset of the other (Baselga, 2010). The number of new OTUs being detected with each new sample indicates that this study was not comprehensive in its description of the diversity of aerial insects in the GRNP (Figure 3-4).

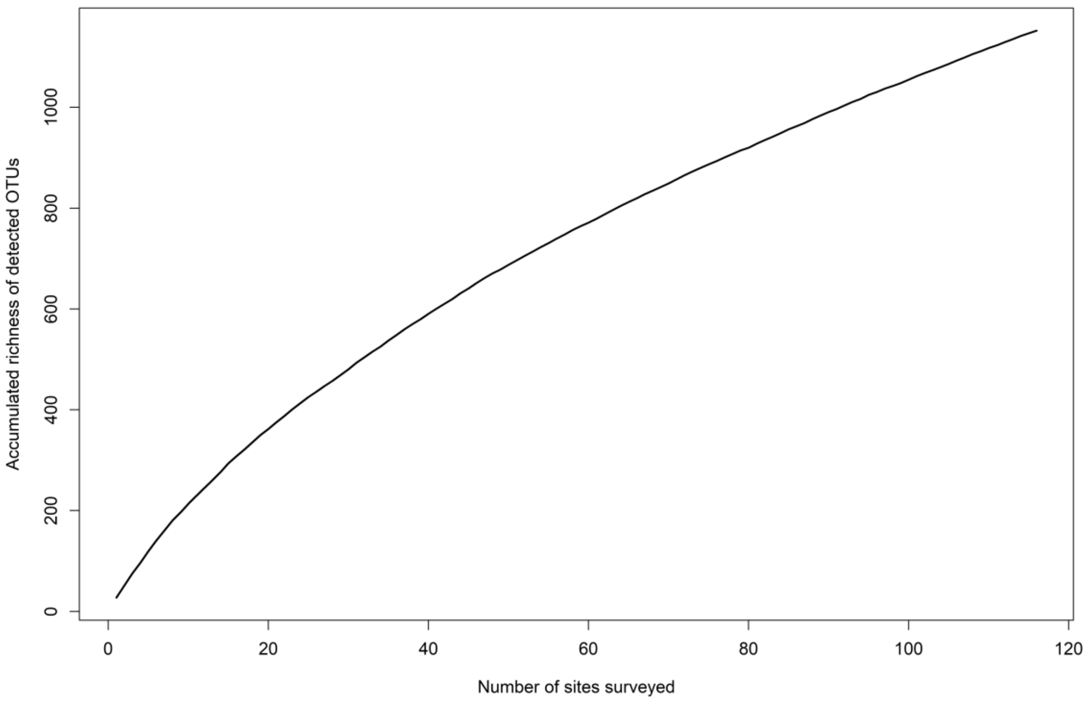

Figure 3-4: The slope of the species accumulation curve indicates that new OTUs were continuing to be found with regularity, and that the full composition of the flying insect community was not close to being detected.

Across the first 360 models comparing satellite platforms and EO variable grouping, the McFadden's pseudo R^2 of proportion of variation explained ranged between 0.19 and 0.39. The ten most effective models were evenly divided among those utilising Landsat-8 (mean R^2=0.385), Sentinel-2 (mean R^2=0.387), and Planet (mean R^2=0.375). The top ten models all incorporated GLCM-Bands, with eight also including GLCM-Indices, while none used raw bands or indices in SCC construction. Given the lack of evidence that the satellite platform used had an effect on results, Landsat-8 was selected as the platform to be used in the ongoing analysis, as its coarser resolution minimised the required computer processing power, and the availability of historic data may facilitate retrospective analysis. Subsequently, using only GLCM-B and GLCM-I to create SCCs, model performance improved when being fit with increasing numbers of between 1 and 20 SCCs, but improvements became negligible when the number of SCCs exceeded 10. Consequently, all subsequent EO testing used 10 GLCM-B and GLCM-I derived SCCs. The buffer defining the area over a which a single site was described produced the best model performance at 30m, with minimal variation observed up to 150m. However, beyond this distance, performances substantially declined (Figure 3-5).

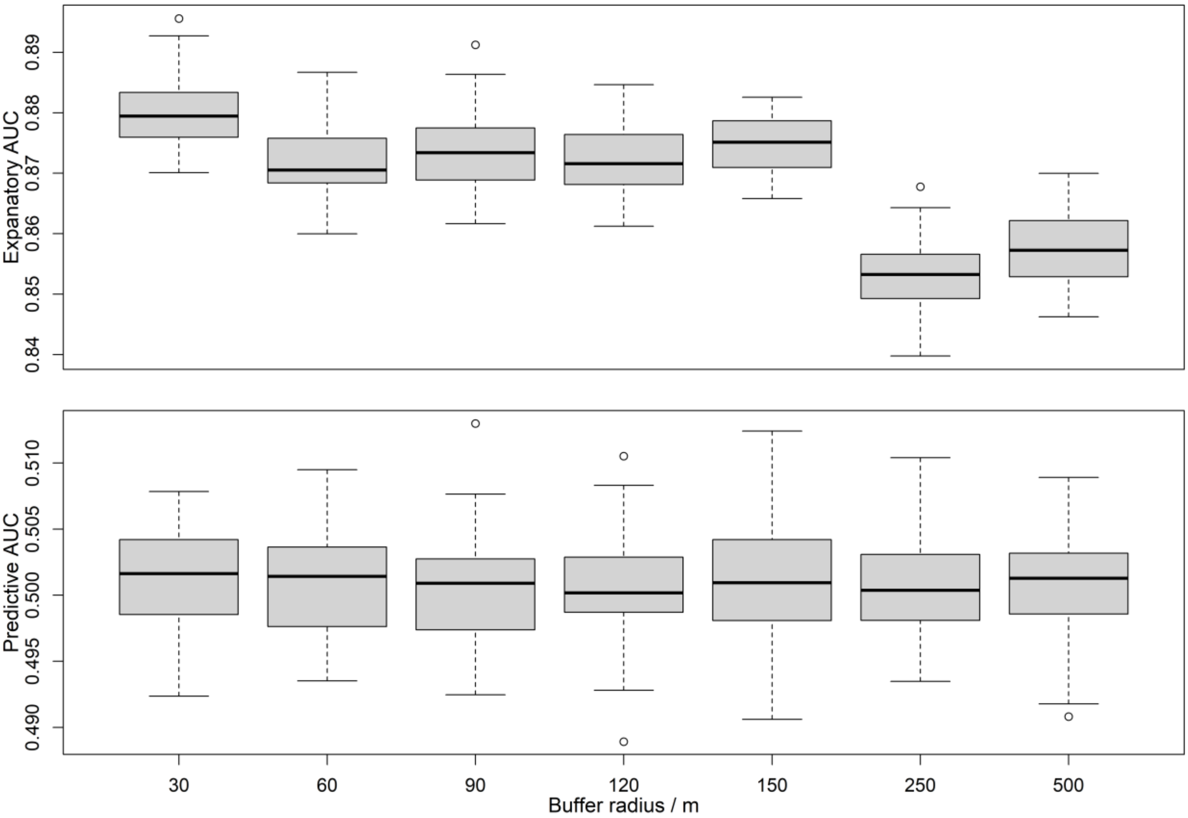

Figure 3-5: The range of AUC of 120 models per buffer size used in measuring reflectance values to create 10SCC's from Landsat-8 data. Top panel shows explanatory AUC, bottom panel shows predictive AUC.

The final full EO sJSDM, fitted with 10SCCs, the year of deforestation, distance to the GRNP, and a radar-derived SCC demonstrated that environmental factors reduced the residual deviance of the null model by 19.5%, whereas co-occurrence among OTUs explained 24.4% of the variation. Spatial autocorrelation explained just 1.8%. The proportion of variance explained is illustrated in Figure 3-6. In predictive models which held out data for validation, the mean explanatory AUC was 0.94, while the mean predictive AUC was 0.58. These values were achieved across all OTUs, with 55 OTUs exhibiting a predictive AUC of at least 0.7 (high-AUC OTUs). By comparison, the model fit with only the habitat variables measured *in situ* showed the habitat variables reduced residual deviance of the null model by 18.0%, accompanied by an increase in the importance of co-occurrence (31.3%) and spatial autocorrelation (9.3%). The predictive sJSDMs fitted to field-data exhibited a mean explanatory AUC of 0.85, and a predictive AUC of 0.56. Of the 49 OTUs with an AUC of at least 0.7, only 12 were common to the high-AUC ETOs described by EO model.

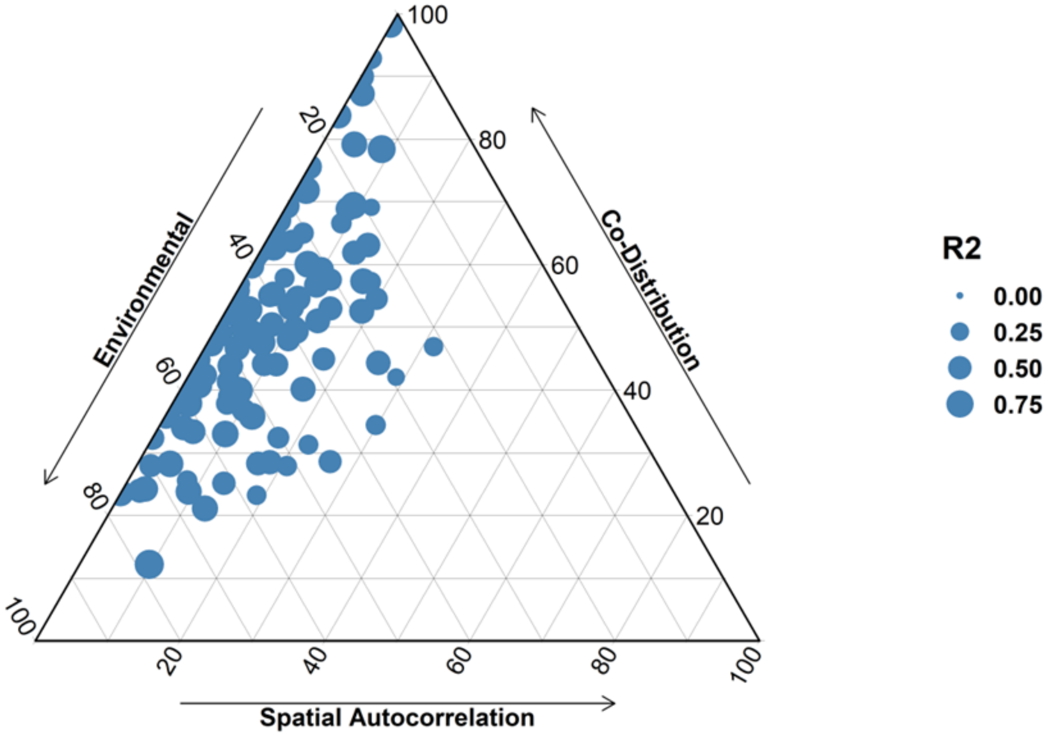

Figure 3-6: The proportional influence of environmental variables, spatial autocorrelation, and OTU co-distribution on site level McFadden's R^2. Each point represents a trap site.

Due to the incomplete nature of community observation, we considered there was a possibility that the high-AUC values were merely a random selection from a normal distribution. To explore this further, we kept the same count values but randomised their locations. For each of four sets of randomisations we re-fitted a model and tested its predictive ability. Using this method, we found high predictive abilities for just 21 OTUs (on contrast to 55 using true count locations). A two-sample t-test showed that the mean of all AUC using the true count data (0.58) was significantly higher than the mean of all AUC using randomised data (0.50) (t(1526) = -8.81, p<0.001). Similarly, as shown in Figure 3-7, a two sample Kolmogorov-Smirnov test showed that the two distributions were significantly different (D=0.211, p<0.001). These results indicate that our high-AUC OTUs were more than just chance predictions.

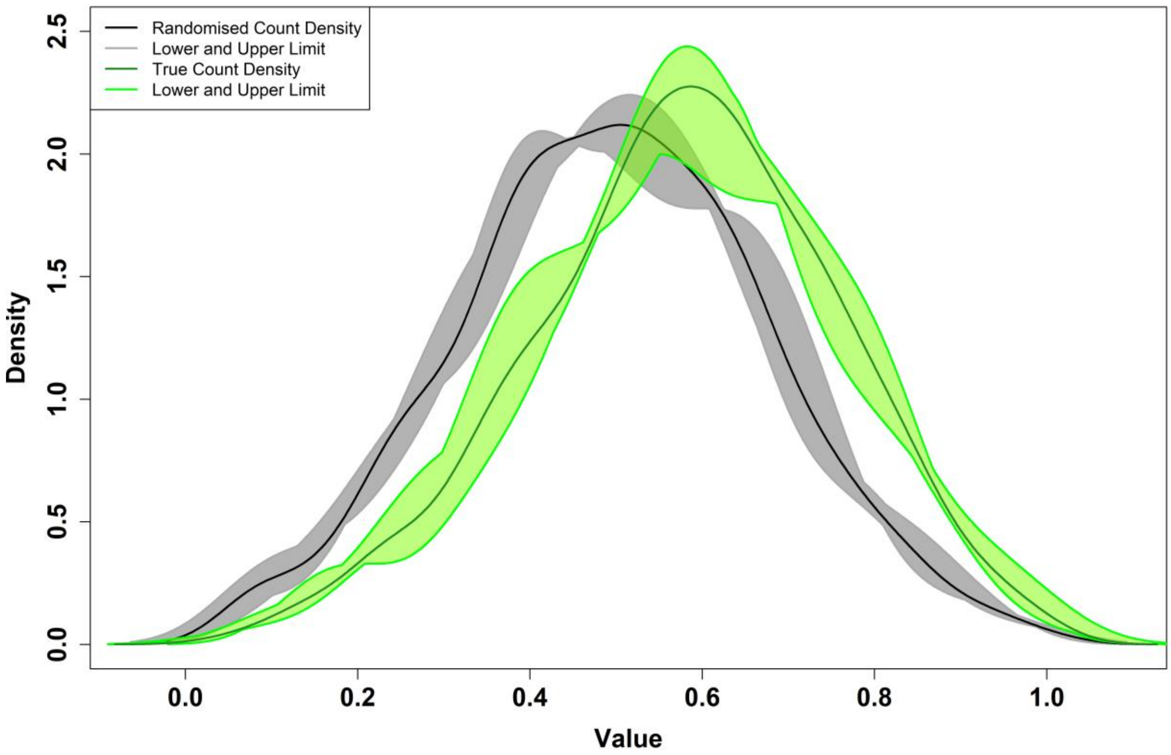

Figure 3-7: Comparison of the density of all predictive AUC values for models run using true and randomised locations of count data.

The variables used in the final EO sJSDM were also found to be capable of explaining a third of the variation in forest structure when modelled with Random Forest (Table 3-3). The three most influential variables and the proportion of variance of maximum canopy height, percentage canopy cover and canopy height diversity explained respectively were: Distance to forest (0.38, 0.26, 0.25); Year of deforestation (0.10, 0.19, 0.17); SCC7 (0.09, 0.11, 0.16).
When predicting insect communities to locations with GEDI information, binomial tests indicated a significant probability that the high-AUC OTUs occurred less frequently with increasing values of forest canopy height, cover, and height diversity (Figure 3-8: Predicted probability of occurrence (y-axes) for the 55 high-AUC OTUs as a function of changes in a) canopy height, b) canopy cover, and c) diversity of canopy height as measured by the GEDI platform. Each line is the line of best fit for an OTU. Red lines show negative associations, and blue lines show positive associations.).

Table 3-3: The mean and standard deviation of three GEDI measured descriptors of forest structure at 2,448 points in the GRNP buffer zone, and the percentage of their variance explained by, and the RMSE of out-of-bag values predicted by random forest models using variables used in the best performing sJSDM.

	Observed Values		Predicted Values	
Forest Structure	Mean measured value	Standard deviation of measured value	Percent of variation explained	RMSE of modelled out-of-bag values
Maximum canopy height	20.9m	13.0m	31.3%	10.9m
Percentage canopy cover	57.7%	32.5%	34.4%	26.6%
Canopy height diversity	2.61	0.62	34.0%	0.51

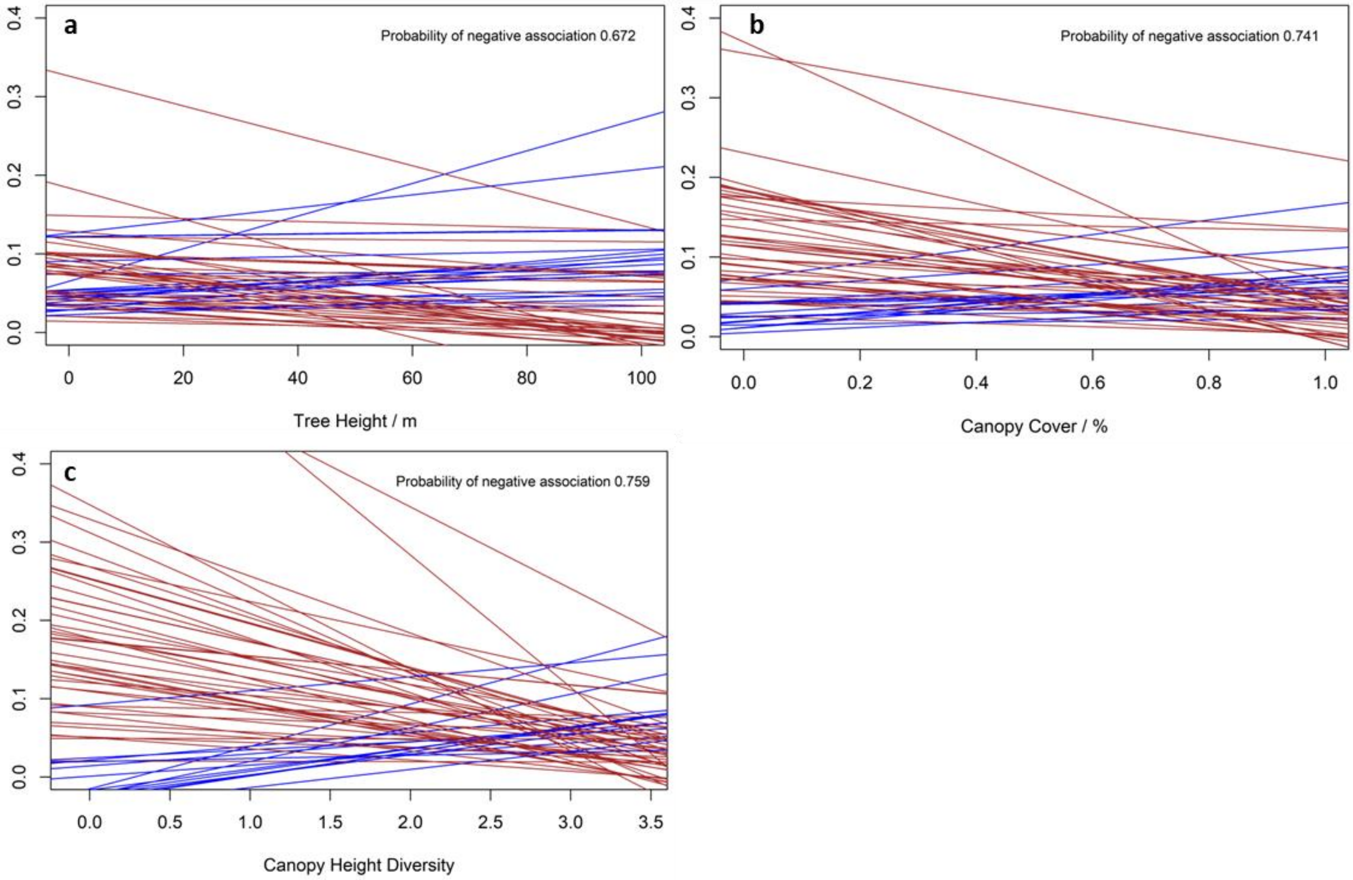

Figure 3-8: Predicted probability of occurrence (y-axes) for the 55 high-AUC OTUs as a function of changes in a) canopy height, b) canopy cover, and c) diversity of canopy height as measured by the GEDI platform. Each line is the line of best fit for an OTU. Red lines show negative associations, and blue lines show positive associations.

The probability of occurrence of the 55 high-AUC OTUs was predicted across a section of the northern and central GRNP and its buffer zone, and community similarities were calculated and mapped (Figure 3-9). The difference in the predicted richness of OTUs with positive and negative correlations with forest structure was calculated and mapped across the central and northern GRNP and their buffer zones (Figure 3-11). A greater richness of OTUs with negative correlations to each of the three descriptors of forest structure was predicted in the buffer zone than inside the GRNP (Table 3-4). These differential richness levels may be utilised as a tool to infer or measure the extent of degradation and environmental health.

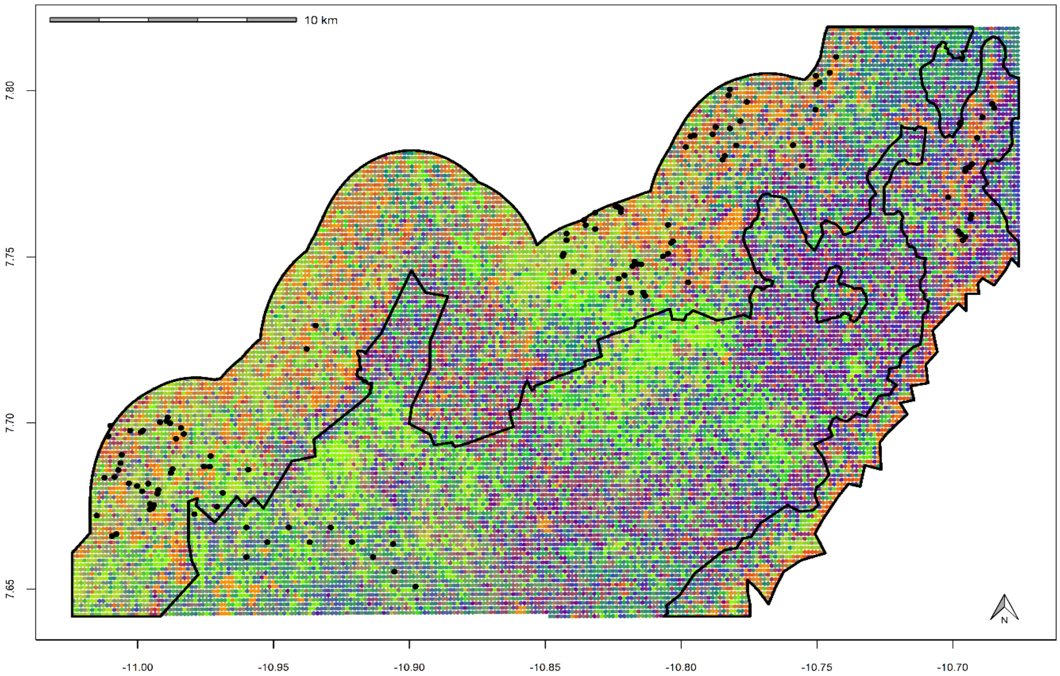

Figure 3-9: Community similarity, based on the 55 high-AUC OTUs, across a section of the northern and central GRNP and its buffer zone. More similar colours represent more similar communities. A guide to colour relatedness is shown in Figure 3-10. Whilst there are observable overlaps of community composition between the buffer zone and main park, orange communities seem more associated with the buffer and purple communities with the main park.

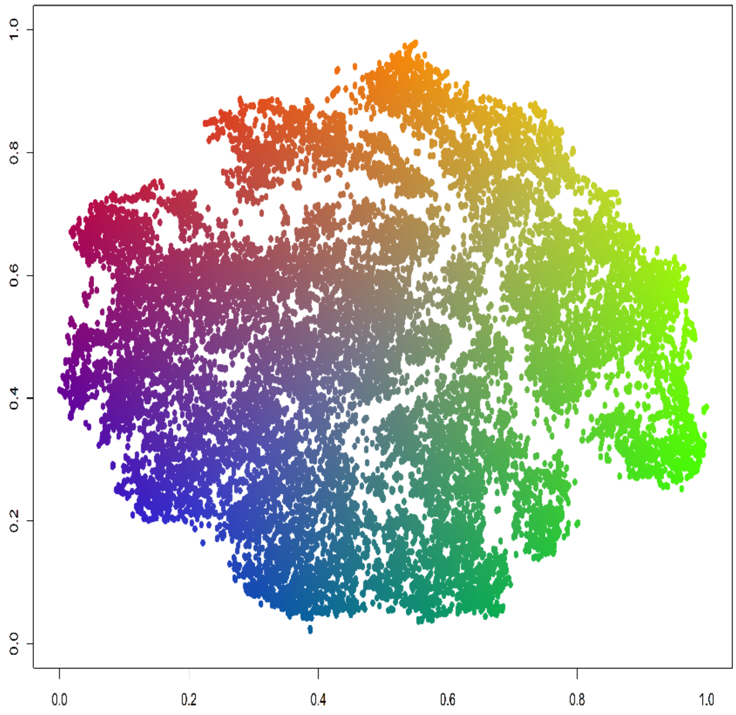

Figure 3-10: Legend indicating community similarity of Figure 3-9 by the represented colour. A difference of 1 indicates entirely different communities and a difference of 0 indicates identical communities.

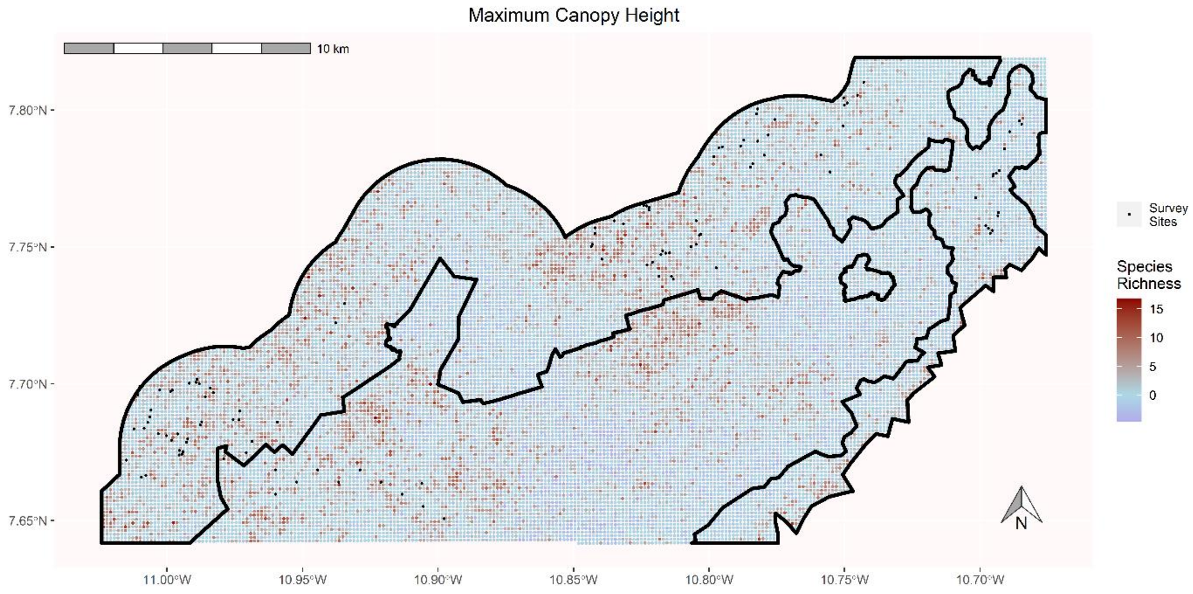

Figure 3-11: The difference in predicted richness of OTUs with a positive correlation to Maximum Canopy Height subtracted from the predicted richness of OTUs with a negative correlation. Higher total richness indicates a greater number of OTUs with a negative compared to positive correlation, and thus potentially indicates where areas have a lower canopy height.

Table 3-4: The mean and median differences in richness between OTUs with significant positive and negative correlations with forest structure, and how they differ between the buffer zone and the GRNP. Higher richness's indicate areas with a greater potential for being degraded as these are areas where OTUs with negative correlations are more prevalent than those with positive correlations.

	Maximum Canopy Height	Proportion Canopy Cover	Canopy Height Diversity	Total Richness
Mean Richness Inside GRNP	0.74	1.02	0.61	4.55
Median Richness Inside GRNP	0.25	0.20	0.01	3.04
Mean Richness Outside GRNP	1.11	3.41	2.00	5.88
Median Richness Outside GRNP	0.40	1.70	0.62	3.83

3.5 Discussion

Effective tropical forest conservation requires not only protecting tree cover, but also conserving biodiversity, including both plant communities and the animal species they support. In forests under anthropogenic pressure, changes in structure can influence these biodiversity patterns. Our results show that remotely sensed structural variation provides meaningful insight into biodiversity across the GRNP buffer zone, supporting the development of monitoring approaches suited to conservation planning. By using insects as indicators of biodiversity, we demonstrated that satellite imagery can predict insect diversity across a tropical forest with a degree of accuracy comparable to, if not slightly superior to, habitat descriptors collected in the field. Although neither dataset was able to reliably predict more than 20% of the OTUs tested, the JSDM fitted to EO data allowed their turnover to be mapped at large scales. Notably, the inverse correlation between many of the well-predicted taxa and structural properties such as canopy cover and height suggests that the subset of taxa highlighted by this study is likely to be more strongly associated with open and degraded habitats than to represent the diversity of insects in closed forests.

3.5.1 Influence of Satellite, Scale and EO variables

Despite the importance of spectral and spatial resolution of satellite imagery in describing the turnover of function and physiology of vegetation communities (Helfenstein et al., 2022), the optical satellite platform from which we extracted spectral data did not influence model performance. The performance of our model was not contingent on a specific optical imaging platform, as the objective was not to identify or classify an individual pixel. The objective was rather to describe how insect species that are invisible to the satellite, are responding to variations in general conditions of the visible surface. Similarly, the model's lack of sensitivity to modest changes in the buffer areas used to calculate spectral means, indicates that local habitats are delineated across numerous pixels rather than being confined to the specific pixel of the trap site. A second and related pattern was the consistent inclusion of GLCM rather than the raw values used to create them. This indicates that the changes being captured by the current model were not responding to mean spectral variation, but instead were responding to changes in spatial patterns that relate to vegetation types (Dorigo et al., 2012).

The presence of cloud cover over tropical forests can make the acquisition of temporally relevant, images, with sufficient clarity, a challenging endeavour (Quiñones et al., 2007). The utilisation of a greater temporal resolution, through the implementation of more frequent satellite passes, can facilitate the generation of a greater number of images from which a cloud-free composite can be constructed. However, the cloud-free composites created using Landsat-8 or Sentinel-2 were of a noticeably inferior quality to the single (almost) cloud-free image obtained for each. Although the pre-processed composite provided by Planet appeared to be relatively clear, a slight image-wide haze was discernible, which may have contributed to the marginally inferior model results obtained using Planet data.

In contrast to passive reflectance, radar imagery is not affected by cloud cover and has the capacity to penetrate the forest canopy to a limited extent, thereby enabling the description of the underlying structure. The combination of radar imagery with spectral imagery has been shown to improve land classification in diverse and dense forests, particularly when textural GLCM measures of radar were included (Mishra et al., 2019; Numbisi et al., 2019). Our results were similarly improved by the incorporation of radar data, but despite our use of GLCM in describing optical reflectance values, we did not follow the same protocol for the Sentinel-1 radar imagery but would do so in future studies.

Radar imagery does have some limitations, as shorter wavelength radar in particular, can have its ability to penetrate vegetation layers impeded by a high-water content in the forest canopy, which reduces the influence of understory vegetation on the reflectance (Cagnina et al., 2023). Thus, variability in water content in or on the leaves due to rain, dew, or seasonal change can cause temporal or spatial variation in radar backscatter for areas of similar cover. The C-band (medium wavelength) radar of Sentinel-1 used in this study is less affected, but is still sensitive to water in vegetation layers, and while it has been shown to be adequate in predicting vegetation indices, it is least effective in forested areas (Lasko, 2022). A new satellite platform, designated NISAR (NASA-ISRO Synthetic Aperture Radar), is scheduled for launch in 2024. It contains a longer wavelength radar which is less affected by surface moisture and better able to penetrate forest canopies and describe sub-canopy structures. Further investigation of both current and upcoming radar imagery may prove beneficial for studies conducted in high cloud areas, particularly when considering phenology, or when temporal description is required congruent with biodiversity surveys.

3.5.2 Insect Communities and Forest Structure

As is common in species rich tropical ecosystems, only a limited portion of the insect community was sampled, and community composition showed complete turnover between sites. This is an expected outcome when sampling mobile, hyper-diverse taxa across heterogeneous landscapes (Feeley and Silman, 2011). Despite this, species co-occurrence explained a substantial portion of variation, showing that species did not associate randomly. Variance partitioning also indicated that descriptors of forest structure had a similarly strong influence on community composition, demonstrating that ecological processes structured these communities in predictable ways. The ability to predict significantly more OTUs than expected by chance reinforces that the data captured meaningful ecological signal, despite being incomplete (Moudrý and Šímová, 2012). This is further supported by the use of DNA metabarcoding, which, while unable to capture a fully exhaustive species list, is well-suited to efficiently sampling broad taxonomic diversity in tropical ecosystems (Bohmann et al., 2014; Yu et al., 2012). These results show that the biodiversity data, while limited, were fit for purpose in assessing broad patterns in the response of aerial insect communities to forest structure.

It is unsurprising that the species accumulation curve indicates, in line with previous studies (Basset et al., 2012; Erwin, 1982; Stork et al., 2015), that the true diversity of arthropods is likely to be much higher than observed. Indeed, Basset et al. (2012) estimated that up to 44,000 arthropod species could exist in a Panamanian tropical forest, with each hectare of that forest containing almost two-thirds of all species that occurred within it. However, the Basset study was unusually intensive, requiring such sustained sampling that it is rarely replicated. Conversely, our study was designed as a pilot to trial EO and collected a relatively limited number of samples. The wealth of insect biodiversity may appear daunting to monitor, but it plays an indispensable role in the overall functioning and integrity of forests. Advances in technology, particularly DNA metabarcoding, have significantly reduced the barriers to studying such diverse taxa, but further work is needed to understand the turnover of the wider community (Zhang et al., 2016).

Wherever studies have had the resources to test biodiversity patterns, they have observed a strong correlation between the composition of plant and insect communities. The anticipated correlation between EO data and forest structure is corroborated by the fact that the same EO variables that are most effective in predicting insect communities could also explain approximately one-third of the variation in forest structural properties. The negative associations between many well-predicted OTUs and GEDI estimates of forest height, canopy cover, and structural heterogeneity suggest a preference for lower and less diverse canopy structure, which is indicative of younger, secondary, or degraded primary forest (Clark et al., 2021). Likewise, a reduction in canopy cover at the local level would indicate a degree of deforestation. Thus, rather than identifying indicators of different forest types, many of the well-predicted OTUs were associated with elements of forest degradation. Only 12 high-AUC OTUs were shared between the EO- and field-models, indicating that each model may be driven by different proxies of degradation. Some i*n situ* descriptions of canopy cover may be identifiable from EO, but other features such as the presence of lianas or shade-cocoa are aspects of the understory that may not be captured by parallel changes in the canopy reflectance. The greater mean richness of degradation-correlated OTUs in the buffer zone also suggests that our model was more adept at predicting the taxa associated with more fragmented or edge habitats. It can be reasonably assumed that these taxa would be less common in the relatively intact NP.

3.5.3 Practical utility for conservation monitoring

The GRNP buffer zone supports a mosaic of land uses, including logging, subsistence farming, agroforestry cocoa, and low-intensity oil palm plantations, all of which influence the effectiveness of the protected area (Laurance et al., 2014). Click or tap here to enter text. Conservation incentive schemes, such as payments for biodiversity conservation can offset this by extending the sphere of influence of protected lands, provided they balance local income generation with biodiversity outcomes (McDonald et al., 2018). Click or tap here to enter text. However, these schemes rely heavily on the ability to monitor environmental condition across the buffer zone, particularly as land use shifts and intensifies. At the same time, demand for large-scale biodiversity monitoring is growing in the private sector, driven by the new Corporate Sustainability Reporting Directive (CSRD), which requires companies operating in the EU to report annually on the biodiversity impacts of their supply chains, particularly in sensitive areas (European Commission, 2022). With thousands of companies affected, the need for scalable biodiversity monitoring tools is set to increase, reinforcing the practical relevance of approaches like the one tested here.

This study demonstrates that EO proxies of forest structure can provide a practical, spatially explicit indicator of biodiversity patterns, with clear relevance to both conservation and land-use planning. The ability to generate biodiversity relevant maps across the landscape means these methods can be used to emphasise where field surveys are most needed. However, this approach is not intended to replace field surveys. Direct biodiversity sampling remains essential for species-level identification, and for capturing fine-scale habitat associations and ecological processes (Yu et al., 2012). Rather than replacing fieldwork, remote sensing offers a spatially scalable complement, identifying where structural change is occurring and highlighting areas that deserve closer ecological inspection. Combining remotely sensed structure with field surveys for ground truthing offers the most effective strategy for large spatial and temporal scale biodiversity monitoring.

The results also highlight some important practical caveats. The extremely high level of species turnover between sites, although a recognised feature in tropical forests (Feeley and Silman, 2011), limits the strength of any single model, reinforcing the need for ongoing field-based sampling. Additionally, although community similarity and richness both provide useful biodiversity proxies, they are not fully independent metrics as similarity is partly shaped by richness differences (Baselga, 2010). This means that mapping richness and similarity separately could be misleading if interpreted without considering the underlying relationship between the two. Nevertheless, used together, these metrics can describe spatial biodiversity patterns across complex, multi-use landscapes that characterise buffer zones.

Finally, although remotely sensed structural data are increasingly accessible through global platforms, the most readily available products have important limitations when applied to biodiversity monitoring at local scales. These off-the-shelf products are often coarse in resolution, updated infrequently, and simplified into broad categories, which mask the finer structural variation that biodiversity responds to. As this study demonstrates, it is possible to generate more ecologically meaningful structural metrics that better reflect the continuous variation in habitat condition across the buffer zone. This type of tailored analysis, linking spectral heterogeneity directly to biodiversity patterns, offers far greater value for conservation management than simply relying on generic global products. If conservation planners want spatially explicit biodiversity relevant information, particularly in heterogeneous landscapes like tropical buffer zones, investing in customised remote sensing workflows like the approach tested here, provides a practical path to more reliable and actionable monitoring data (Skidmore et al., 2015). Overall, this study illustrates that, even with limited biodiversity data, EO derived forest structure can provide meaningful insights into spatial patterns of biodiversity across tropical buffer zones.
Click or tap here to enter text.

3.5.4 Conclusion

There is a substantial body of empirical evidence supporting the assumption that insect diversity is strongly associated with vegetation structure and diversity. However, the mechanisms underpinning this association are not easily resolved. Moreover, studies have encountered difficulties in describing the diversity of plant communities within tropical environments from satellite imagery due to the extremely high plant diversity and lack of change in reflectance across seasons (Rocchini et al., 2016; Torresani et al., 2019). The high plant and insect diversity observed in tropical forests, coupled with the high turnover observed among Malaise samples, suggests that local and landscape-scale estimates may be substantially underestimated (Basset et al., 2012). Despite these challenges, this study demonstrated that EO derived metrics explained variation in insect community composition across the buffer zone, with predictive power varying widely between taxa. While some taxa were predicted particularly well, the overall signal across the community was modest but clearly non-random, indicating that structural proxies captured ecologically meaningful variation in habitat condition. This is encouraging given the limited sampling conducted, and it suggests that with more comprehensive field data to calibrate models, a significantly higher proportion of taxa could be predicted using similar approaches. Furthermore, it is possible that relationships identified could become more pronounced if more sophisticated sources of remote sensing become available and were included in future work.

4. Assessing variations in Peruvian Amazon bird communities and the utility of Landsat variables for biodiversity monitoring.

4.1 Abstract

The extent of deforestation and degradation of the Amazon rainforest is significant, with more than a third of the forest lost or altered as a result of various human activities, including logging, agricultural expansion, and fires. While the role of Earth observation satellites in delivering and reporting on the success of policies to preserve forest cover is widely appreciated, the impact of site and landscape degradation on the distribution of biodiversity within the remaining forests has proven far harder to assess at large scales. This study concentrated on the Tambopata forest in the Madre de Dios region of south-eastern Peru which has a long history of monitoring, to explore whether bird communities respond in line with changes in habitat type and indicators of forest degradation. A total of 3129 surveys were conducted over a 16-year period, and the drivers of community turnover were examined along with the power of existing monitoring approaches to detect potential future change. A high-dimensional description of the forest canopy and land surface was created using Landsat imagery, being the platform which provided the highest resolution imagery over the period of study. We tested the potential of both habitat descriptors and Landsat derived remotely sensed (RS) variables to predict the probability of both occurrence and detection of the 135 species most frequently observed of 358 species identified in total. Models based on Landsat reflectance had superior predictive performance for bird species occupancy (mean AUC = 0.68), compared to models trained on habitat data (mean AUC = 0.58). Furthermore, they exhibited high predictive ability (AUC>0.7) for a greater number of individual species (49 compared to 20 for each data set respectively). Species detection rates were found to be very low, meaning the true richness of bird communities was possibly greater, and the variation in community composition across the landscape lower than that of the communities observed. Consequently, predictions of community composition were found to be less effective than those of individual species.

Despite the promising potential of remotely sensed data in species distribution models, the low detection rates of many species significantly reduce the ability of models to accurately estimate community composition and detect changes in community structure over time. To increase detection rates and thus the ability of models to better predict communities, we propose survey designs that concentrate on more replicates at fewer sites. Should conservation projects set clear trigger points, such as predefined percentage increases or decreases in species occupancy, surveys could be designed to ensure a defined probability to detect those changes. Our findings highlight the potential for satellite remote sensing to enhance biodiversity monitoring beyond the measurement of deforestation and land use. By integrating refined survey design with RS data, the detection of species declines and habitat degradation can be improved, thereby strengthening adaptive management strategies for the conservation of tropical forests.

4.2 Introduction

Tropical forests are the most biologically diverse areas on Earth (Bradshaw et al., 2009; Gibson et al., 2011), and more than a billion people depend directly on them for food, materials, and services (Lewis et al., 2015). Tropical forests also play a key role in the global carbon cycle, accounting for approximately half of terrestrial carbon uptake and accounting for up to two-thirds of the carbon sink in biomass, soil, and deadwood, of all forests worldwide (Hubau et al., 2020; Pan et al., 2011). However, clear-cut logging and commercial agriculture are altering or completely destroying forest habitats. This is evidenced by the fact that over 17% of the Amazon Basin, the world’s largest continuous tropical forest, has been deforested and 38% of the remaining area has been degraded (Lapola et al., 2023; Potapov et al., 2017; Vergara et al., 2022).

Conservation strategies, such as the Reducing Emissions from Deforestation and Degradation (REDD+) initiative, are believed to have had a generally positive, albeit variable, influence in reducing deforestation, enhancing carbon stocks, and improving sustainable forest management (Guizar-Coutiño et al., 2022; UNFCCC, 2023; Wunder et al., 2024). However, despite reductions in deforestation and increases in the amount of tropical forest falling under protection, half of the world’s tropical forest reserves continue to lose taxonomic and functional biodiversity due to anthropogenic disturbances other than forest cover loss (Barlow et al., 2016; Laurance et al., 2012). The loss of forest cover can be used to estimate the level and location of carbon loss and is routinely mapped in near real-time (Hansen et al., 2013), but to assess the efficacy of biodiversity conservation strategies, such as REDD+, beyond forest cover conservation, it is necessary to ascertain the current condition of biodiversity within the ecosystem and where and how it changes in the future (Stork et al., 1997; Titeux et al., 2016; Willis et al., 2007).

The characteristics of ecosystems can be described structurally by the type, quantity and distribution of elements within them (Pettorelli and Schulte to Bühne, 2022). However, categorical elements such as whether or not an area is a forest, is intact, has been logged, is secondary, is a plantation, can be ambiguous or dependant on viewpoint, and require prior knowledge to define them (Chazdon et al., 2016; Savilaakso et al., 2023). Despite any *a priori* assumptions, the selected elements may not be the most effective descriptors of an ecosystem’s condition. In this regard, essential biodiversity variables, including species distribution, abundance, and community composition, have been proposed as additional criteria to monitor (GEO BON, 2018; Pereira et al., 2013). Thus, defining the ecological health of forest environments and identifying where and how biodiversity is changing requires monitoring and description that goes beyond simply quantifying deforestation levels. Satellite remote sensing (RS) cannot directly see essential biodiversity variables, but it allows for the measurement of deforestation along with land cover and habitat structure (Hansen et al., 2013). RS can also provide a broader, scalable environmental descriptors, alternative to field surveys, that act as proxies for essential biodiversity variables, for example, ecological diversity, phenology and physiology (Asner et al., 2013; Pettorelli et al., 2005; Purdon et al., 2022; Zhang et al., 2003). Continuous remotely sensed variables have also been shown to be better than land class at describing plant species richness and diversity (Perrone et al., 2023). RS surface reflectance values can therefore be used to infer vegetation and environmental structure as continuous gradients, rather than discrete categories, over a region of interest. This approach avoids the need to pre-define habitat classes that may overlook important ecological variation and reduces the mismatch between habitat classification and the scale at which they can be measured. While RS variables can infer potential structure and ecological richness, there is little understanding of how they relate to ecological communities within forests (Ludwig et al., 2016; Pettorelli et al., 2014b; Turner et al., 2003).

Thus, the measurement of structure or deforestation alone provides only a partial description of an environment's condition, but faunal biodiversity can regulate forest dynamics (Barlow et al., 2016) and be used as an indicator of ecosystem state (Dirzo et al., 2014; Gerlach et al., 2013). Many studies have shown that land use and vegetation structure influence the biodiversity of an area. For example, dung beetle community composition and soil bioturbation are influenced by the intensity of logging within tropical forests and the spatial scale at which it is measured (França et al., 2017). Similarly, the richness and biological function of mammal communities are impacted by proportional levels of deforestation and the resulting size of forest patches in Atlantic Amazonian forests (Magioli et al., 2021). Furthermore, insect community richness and composition have been shown to respond to and reflect forest structure, forest composition, plant diversity and forest age (Basset et al., 2012; Schowalter, 2017; Traylor et al., 2022; Zhang et al., 2016). Finally, the structure of bird communities has been used both to assess the effectiveness of reforestation (Barros et al., 2022), and to show that they are negatively affected by patch size reduction within the Brazilian Amazon (Bregman et al., 2015).

However, the biodiversity surveys required to adequately monitor indicator taxa are time consuming and expensive, thus limiting the effective area that can be monitored in space and time (Rocchini et al., 2016; Skidmore et al., 2015). Nevertheless, results from biodiversity surveys can be linked in statistical models with RS descriptors of the Earth's surface, to estimate community and ecosystem structure across a landscape with limited expense (Csillik et al., 2019). We would expect Amazonian bird communities to be broadly influenced by the characteristics of the forest in which they occur, and if changes in forest structure led to changes in bird communities, we would expect that remotely sensed surface reflectance values to correlate with this.

The aim of this study was to test whether RS data can be used to predict changes in the diversity of bird communities across the forests of the Tambopata Forest of south-eastern Peru. We hypothesise that bird community composition will vary predictably between habitat classes, and that continuous RS derived surface reflectance variables will explain a significant proportion of this variation. Furthermore, we expect that the predictive power of continuous RS descriptors will exceed that of categorical habitat descriptors. If successful, predicted bird communities could act as an indicator of changes in forest biodiversity and condition. We first test whether broad categorical habitat descriptors, such as landscape type, influence the bird species observed, and assess how well these variables predict community composition. We then ask whether continuous RS surface reflectance data can predict the same bird communities. By using continuous data, we aim to capture gradual changes in vegetation and habitat structure, rather than treating habitats as fixed categories, which may overlook important ecological variation. This approach could provide a more accurate way to describe environmental conditions across the landscape. Furthermore, the continually updated nature of RS data could allow for this method to be used in ongoing monitoring to detect changes in community structure. As such, lastly we explore how the number of survey stations and replicate visits affects the power to detect change, helping to guide efficient survey design. By bringing these elements together, we aim to provide a practical tool for policy makers and ecologists to assess the successes and failures of management strategies, and to support decisions about financial incentives.

4.3 Methods

4.3.1 Study area and surveys

Survey data were collected by FaunaForever (www.faunaforever.org) to monitor species diversity across an array of landscapes in the Madre de Dios region in south-eastern Peru, close to the Tambopata River south-west of Puerto Maldonado (Figure 4-1). Conservation researchers, assisted by volunteers, performed assorted bird surveys across a total of 637 stations between 2004 and 2020. The stations were clustered around 25 centres that were located between 1.5km to 170km apart. Within the centres, stations were typically situated between 200m and several kilometres apart and were each surveyed between one and 35 times (median = four), representing a total of 3,129 surveys. Some stations were surveyed in more than one calendar year, and because populations and forest structure can change over extended periods of time, we treated these as independent stations for each year. This resulted in a total of 1,135 independent stations, which were surveyed between one and 21 times, with a mode of 1 survey, a median of 2 surveys and a mean of 2.9 surveys per site (SD ±2.4). Surveys were conducted as point counts, where all bird species seen or heard by the researcher during ten minutes at each point were recorded, or as net counts, where a 36m long mist-net was deployed and all species caught were recorded. A total of 358 species were identified, but in line with suggestions for minimum observations based on mean prevalence (J van Proosdij et al., 2016), only the 138 species that were observed at more than 10 stations were used in training models for full analysis.

4.3.2 Environmental variables

Habitat was classified using two categories: land class and forest degradation. Land classes were floodplain forest, terra firma forest, or agricultural land, and were determined using the MapBiomas Peru Project, Collection**-1** of the Annual Land Cover and Land Use Series for Peru 2013; (MapBiomas, 2023). Forest degradation was determined from a primary forest raster using the 2000-2019 forest cover loss data set (Hansen et al., 2013). Pixels were classified as primary forest if no deforestation year was given, assuming they had remained intact. Subsequently, secondary forests were defined as areas that had undergone recorded deforestation since 2000. This classification is primary/not primary and does not differentiate between successional stages of secondary forest. The proportion of non-primary forest, equating to forest assumed to have been deforested or degraded, was calculated within three different radii around each station (1km, 2km and 5km). No cross validation was performed between the MapBiomas and Hansen datasets, resulting in possible inconsistencies, where areas classified as forest in MapBiomas may not necessarily align with the forest loss described by Hansen.

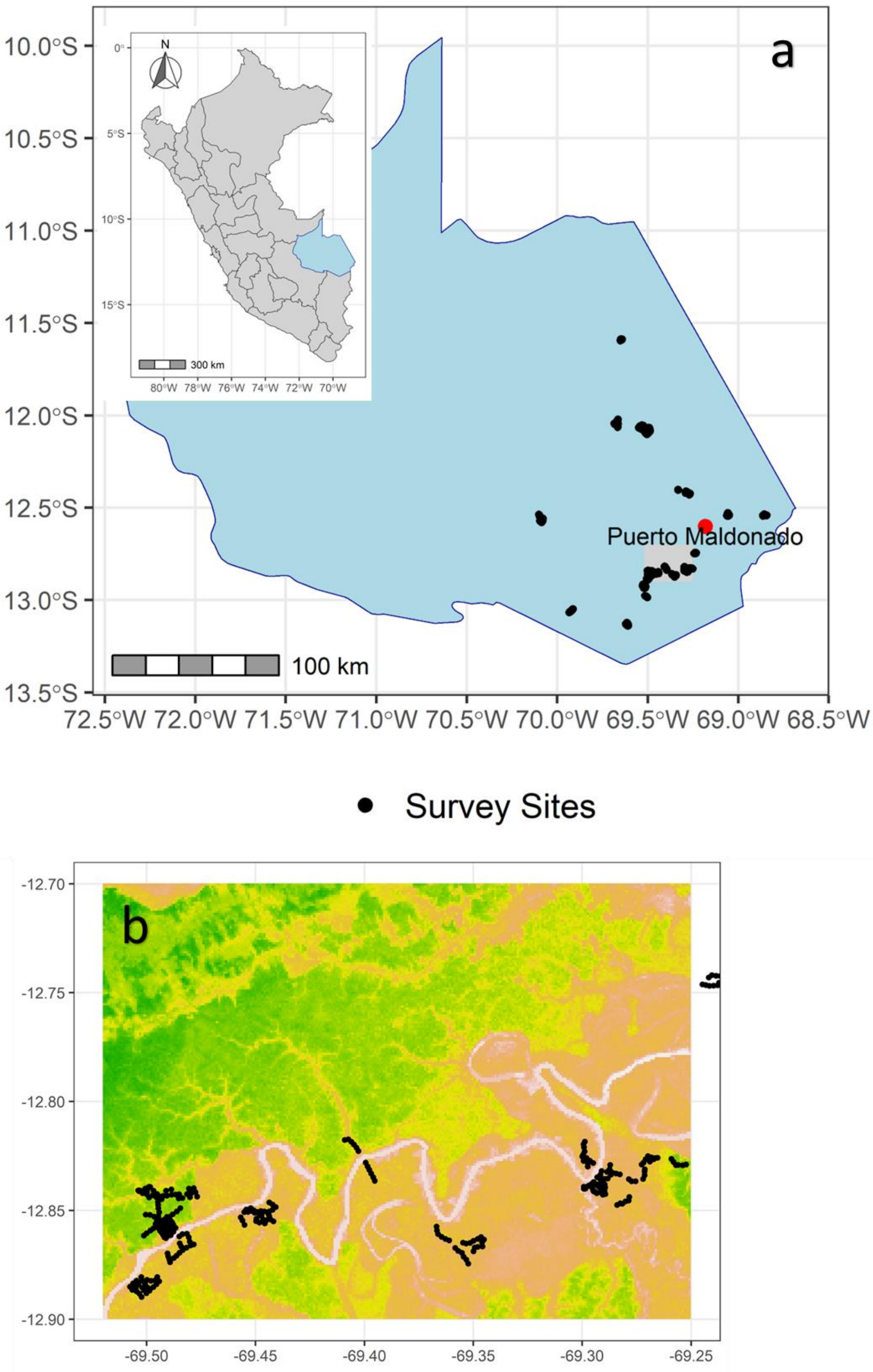

Figure 4-1: (a) The location of survey stations within the Madre de Dios region. Inset is Peru with Madre de Dios highlighted in blue. Panel (b) displays the distribution of survey stations within the grey box highlighted in panel (a) along the Tambopata River within regional topography (Farr et al., 2007).

Of the 1135 surveys, 489 were on terra firma, 603 on floodplain and 43 on agricultural land, and 1109 were defined as being inside primary forest, and 26 in secondary forest. Elevation can influence the distribution of tropical forest bird species (Carvalho et al., 2023; Nores, 2000). However, all survey points were located within the narrow range between 175m and 310m above sea level, so it was deemed unlikely that elevation would significantly affect model performance; it was therefore not included in the analysis.

4.3.3 Satellite derived variables

Reflectance signatures at different wavelengths can be employed to assess the greenness and moisture content of vegetation, with some correlation observed between these signatures and the leaf structure of a plant (Huete, 2012). The combination of reflectance and water content descriptors is a valuable tool for describing the physiology of the forest canopy (Ollinger, 2011). The spectral diversity within a canopy cannot be measured by a single metric (Wang et al., 2018), and as the spectral characteristics and spatial scales that best capture variation in the forest canopy are unknown, a variety of indices were created to describe variation within the canopy. While many satellite platforms now provide passive optical surface reflectance data, as well as active radar reflectance, the Landsat missions were the only platforms that covered the entire period of the field surveys (Landsat-5 for 2004-2011, Landsat-7 for 2012, and Landsat-8 for 2013-2020). Frequent cloud cover over tropical forests can greatly reduce the number of available images that have sufficient clarity for analysis (Quiñones et al., 2007). To compensate for the low visibility of ground within individual images, a cloud-masked, median pixel value image was generated in Google Earth Engine for each year of study. There is a compromise to using annualised composite images, as reflectance changes caused by seasonality such as leaf flushing and fruiting cycles, or logging and burning, may be masked (Morton et al., 2005; Zhang et al., 2003), although phenological variation in tropical forests can be lower than in some other landscapes (Hilker et al., 2014). Each image contained six reflectance bands, which were then used to calculate seven vegetation indices: the normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), normalised burn ratio (NBR), normalised difference water index (NDWI), and the trio of tasselled cap wetness, greenness, and brightness (TCW, TCG, TCB). Both NDVI and EVI were used to detect vegetation cover and vigour. NDVI was included as one of the most commonly used indices, but it is prone to saturation in dense forests such as the Amazon, whereas EVI is better able to penetrate beyond the canopy surface and incorporate sub-layers. We used NBR to highlight freshly opened soil, and NDWI to highlight water content within the canopy. Tasselled cap indices provide an alternative way of describing reflectance and structure, and have been shown to correlate with, but show more detail than, NDVI (Samarawickrama et al., 2017).

For each survey conducted in each year, the mean and standard deviation of all pixels within a 100m radius of the survey station were calculated for the six reflectance bands, and seven indices. The 26 variables collectively provide comprehensive and high-dimensional description of the reflectance of the local vegetation structure. As many of these variables were likely to be redundant we used principal component analysis to reduce the 26 variables to five components that explained 96% of the variation in the data. The distribution of the five principal components and the proportion of deforestation at three scales are plotted in Figure 4-2, and the relationship between the first two principal components with environmental descriptors is shown in Figure 4-3.

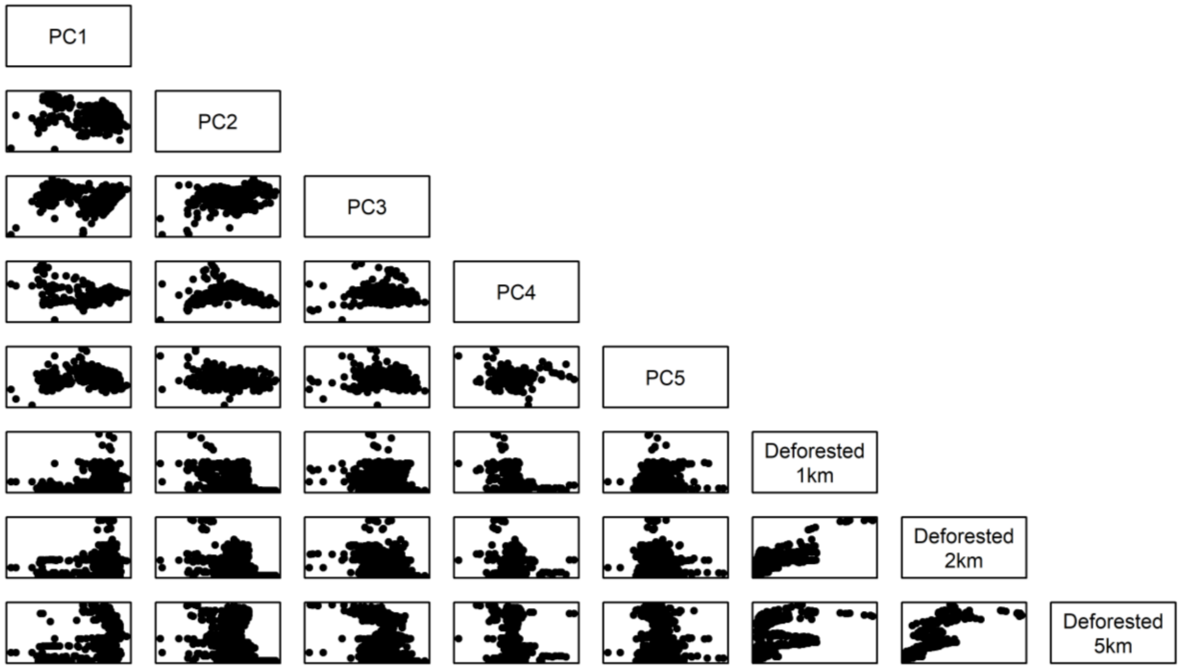

Figure 4-2: The distribution of variables used in EO models, comprising the five EO derived principal components, and the proportion of deforestation calculated at the three measured areas.

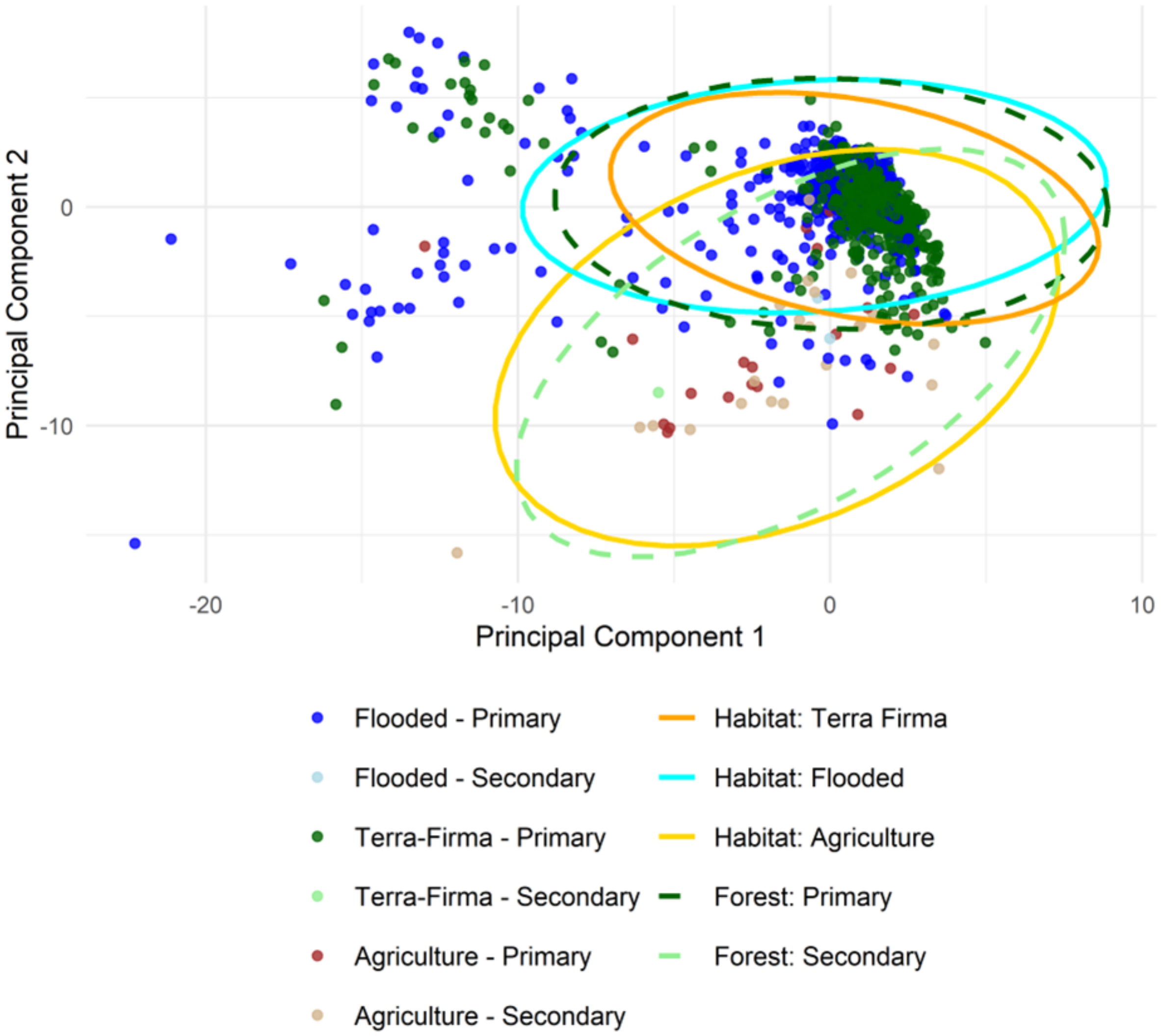

Figure 4-3: The first two principal components, representing 76% of the variation in the RS data, plotted against each other and showing their relationship with the measured habitat types. Each point represents an individual survey and is coloured based on the combination of habitat (floodplain forest, terra-firma forest, agriculture) and forest type (primary or secondary). Ellipses represent the 95% confidence interval around the mean for each categorical group.

4.3.4 Analysis

To quantify the probability of species detection, and to identify whether the composition of bird communities can be predicted as a function of habitat or RS descriptors, we used spatial factor multi-species occupancy models. The probability of a species being observed on a given survey is a combination of its presence (1 = present, 0 = absent; assumed to be fixed over the year of surveillance), and the probability of detection. Detection can be influenced by survey method. For example, mist nets set at ground level will tend to under-sample species which mainly occupy the canopy, as well as birds which are large or less mobile (Remsen, and Good, 1996). Furthermore, the species captured can be influenced by the type of net used (e.g. mesh size), along with the duration of the netting period. Point counts are good for detecting vocal species but may under-sample silent or secretive species, and their effectiveness can be influenced by background noise, weather and observer skill (Rosenstock et al., 2002; Simons et al., 2007). As a result, perfect detection of every species in a community is essentially impossible. The non-detection of a species can bias covariate estimates and predictions of community composition in distribution models (Kery, 2008). However, occupancy models account for imperfect detection by calculating the probability of occurrence of a species at a survey station, and provide probabilities of occupancy at all stations, whether observed or not (MacKenzie et al., 2003). Occupancy models also provide detection probabilities for all species at all stations and the influence of survey specific covariates on these probabilities (Royle and Dorazio, 2009). After the model had been fitted and species detection and occurrence probabilities ascertained, the methods of Guillera-Arroita and Lahoz-Monfort (2012), were used to calculate the power of the study. The term “power” is defined as the probability that a proportional reduction in occupancy of a species will be detected between it and another survey conducted at a different time. We then calculated the survey effort required to achieve a mean power across all species of 0.7, to detect a 50% reduction in their occupancy. A power of 0.8 is frequently used, but with the very low detection rates we experienced, 0.7 was a more feasible target and falls within the test ranges ((Guillera-Arroita & Lahoz-Monfort, 2012). Furthermore, we explored how the power to detect various reductions in occupancy changed with the number of stations and replicate surveys, which in turn could be used to recommend survey regimes for future monitoring.

Four models were compared, one was fitted with habitat variables, one with principal components derived from RS reflectance variables, one with no occurrence variables to act as an intercept-only baseline for comparison and lastly, to compare and examine the influence of spatial factors, the RS model was then re-fitted without spatial factors. All models included survey method (point or net) and time of day (a.m. or p.m.) as covariates for the probability of detection. Although some stations were surveyed up to 21 times, 99% of all stations were surveyed 12 times or less. To improve the efficiency of model fitting and reduce the computational resources required to run them, we fitted the modes using a maximum of 12 replicates per station.

Models were fitted using the Bayesian occupancy model package, “spOccupancy” (Doser et al., 2022). The “spOccupancy” package enables the modelling of all species simultaneously, and accounts for autocorrelation between species. In addition, spatial latent factors, which are akin to unmeasured site-specific covariates that decay with distance, can be included to account for residual species correlations. The number of spatial factors used should adequately explain variation within the community but to maximise efficiency and avoid overfitting, particularly where there are relatively few replicate surveys and species are rarely detected, it is better to include fewer rather than more. We chose to include three spatial factors, each of which is independent and derived from the residual species correlations that remain after correlations with all other variables and previous spatial factors have been calculated. Each spatial factor provides a value (phi) indicating the rate of decay of spatial autocorrelation. The distance at which stations are no longer autocorrelated for a given factor is calculated as 3/phi.

Models were fitted with 50,000 iterations, of which the first 25,000 were discarded as burn-in, and the remainder thinned by 50 to produce 500 posterior samples. Each posterior sample provides a probability of occurrence (psi) for each species at each station and a latent presence/absence (1/0) (z) calculated from this probability. In addition, a probability of detection (given presence) (p) is provided per species per station for a single survey. Chain convergence is particularly difficult in spatial models, so models were fit using a single chain, and mixing assessed visually and using the Geweke diagnostic (Doser, 2023). In fitting the models, we assumed no false presences and treated the entire study as a single closed season. Goodness-of-fit for each model was measured by performing a posterior predictive check to calculate a Bayesian p-value, based on the differences between chi-squared values of observed and predicted data per species, grouped across stations. All fitted models had Bayesian p-values close to 0.5 (0.467-0.477) indicating a good fit to the data (Hooten and Hobbs, 2015; Kéry and Royle, 2016), and models were therefore subsequently compared using the widely applicable information criterion (WAIC) (Watanabe, 2010) and metrics of predictive performance. Performance was evaluated at the species-level, using area under the receiver operating curve (AUC), and at the station community-level, using species accumulation and Bray-Curtis's dissimilarity of community composition.

Despite the known differences in performance of models on training data and withheld datasets, predictive performance on independent data is rarely undertaken with occupancy models (Gould et al., 2019). We did so here by withholding a validation dataset which included two stations from each field centre, each of which had been surveyed at least three times, and in total represented 5% of the dataset. To validate model performance, it was necessary to incorporate both occurrence and detection probabilities when comparing predictions with observations. When predicting species occupancies, each of the model's posterior samples predicts two occupancy values. First is a probability of occupancy (psi) per species per site, from which the second value of latent occupancy (z) of 1 (present) or 0 (absent) is drawn. The probability of detection (p) for each species and site in a single survey is predicted, and this does not vary between posterior samples. The total probability of detection for each species and station is therefore determined by the number of replicate surveys conducted at that station and was calculated as $1 - (1 - p)^n$, where n is the number of replicate surveys performed. This overall detection probability was used to create a binomial detection/non-detection (1 or 0) event per species per station. The latent detection values were then multiplied by the predicted probability of occurrence (psi) and latent occupancy (z) of the same species and station for each posterior sample. Thus, we used the predicted probability that a species was seen rather than occurred when determining model performance against validation data.

4.4 Results

4.4.1 Model fit

All models had Bayesian p-values between 0.467-0.477, indicating a good fit to the data. The model with the lowest WAIC value, thus inferring it was the best performing model, was fitted with RS reflectance data and included spatial factors, followed by the covariate free intercept model, the RS reflectance model fitted without spatial factors, and the highest WAIC indicating the worst performance was from the model fitted with environmental descriptors.

The probability of detection (p) was consistent across all models. Surveys conducted in the morning, using net counts have median species-level p = 0.022, 95% CI [0.001, 0.179]. When conducting surveys in the afternoon, median detection probability significantly lower, at p = 0.015, 95% CI [0.001, 0.018]. When using point rather than net surveys, the median p increased significantly to 0.065 95% CI [0.000, 0.389]. However, point surveys did not significantly increase detection for three of the nine Orders represented. Those three were Caprimulgiformes (nightjars) and Coraciiformes (kingfishers and hornbills) and Galliformes (ground dwelling birds).

The mean probability of occurrence across all species in primary forest and an agricultural habitat, was 0.291, 95% CI [0.211, 0.375]. In secondary forest, the mean occupancy was higher than in primary forest, rising to 0.47, 95% CI [0.36, 0.61]. In comparison to agricultural landscapes, mean occupancy in floodplain increased to 0.36, 95% CI [0.30, 0.47], and in terra-firma mean occupancy increased to 0.37, 95% CI [0.30, 0.47] habitats. Mean occupancy was not shown to vary significantly with deforestation within a 1-2 km radius, but increased by 0.057, 95% CI [0.008, 0.108] with each 10% increase in deforestation within a 5km radius.
When using RS data, the model suggested that spatial autocorrelation extended to 5km, 95% CI [3.3km, 7.5km] for one factor and 7.5km, 95% CI [4.3km, 15km] for each of the other two factors. When using habitat data, the model suggested that spatial autocorrelation extended to 2km, 95% CI [2km, 2.3km] for one factor, to 30km, 95% CI [3.75km, ∞] for another, and the third factor was not significant. These factors represent unmeasured variables that influence the occurrence of species, and fade with distance. Although we do not know what is driving this spatial structure, it can be influenced by anything from local vegetation cover to regional climate or land use. Significant spatial autocorrelation confirm that patterns exist beyond those explained by the known environmental variables.

4.4.2 Model validation

The predictive performances of models followed the implications made by their WAIC values, and the model that produced the best predictive results was also fitted with RS reflectance data and included spatial factors. When assessed on training data, this model had a mean AUC of 0.88 (SD=0.06) and when predicting to independent validation sites, had a mean AUC of 0.68 (SD=0.18). Species for which predictive performance was considered good (AUC>0.7) (Mandrekar, 2010) are hereafter referred to as high-AUC species, and this model had 49 high-AUC species, the most of any model. The performance of all four models is shown in Table 4-1. A species taxonomic order was not shown to have any correlation with its AUC.

Table 4-1: The fit and predictive performance of the four models analysed. Low WAIC values and high AUC values represent better performances. It can be seen that the best performing model in all categories is fitted with RS reflectance data and includes spatial factors.

Model Covariates	Spatial Factor Included	WAIC value	Delta WAIC	Mean Explanatory AUC	Mean Predictive AUC	Number of high-AUC species
RS Reflectance	yes	66773	0	0.88	0.68	49
RS Reflectance	no	68620	1847	0.87	0.62	31
Habitat Descriptors	yes	69432	2659	0.87	0.58	21
None	yes	66995	222	0.8	0.66	43

The predicted species richness of survey stations (6.86 ± 5.1) was close to the observed average richness (6.00 ± 3.6), but when extended to the entire training dataset the model predictions suggested that it would take much longer to detect the total number of species included (Figure 4-4(A)). The mean predicted richness of test stations (9.94 ± 3.8) was also comparable to the observed mean richness (10.82 ± 5.3) and indicated a similar accumulation of species, both across stations (Figure 4-4(B)), and with increasing replication within a station (Figure 4-4(C)). The species accumulation curve in Figure 4-4(C) come from one station that was surveyed 12 times and illustrates the extent to which observed accumulated richness greatly underestimates the total richness expected to be present at each location. A consequence of this was that only a fraction of the species that occur at a station were observed, and therefore the average Bray-Curtis dissimilarity between the list of birds observed and those predicted was 0.8 (± 0.1). In locations with higher observed richness, the surveyed total represented a greater proportion of the true expected total, and the dissimilarity between predicted and observed communities was substantially lower (Figure 4-5).

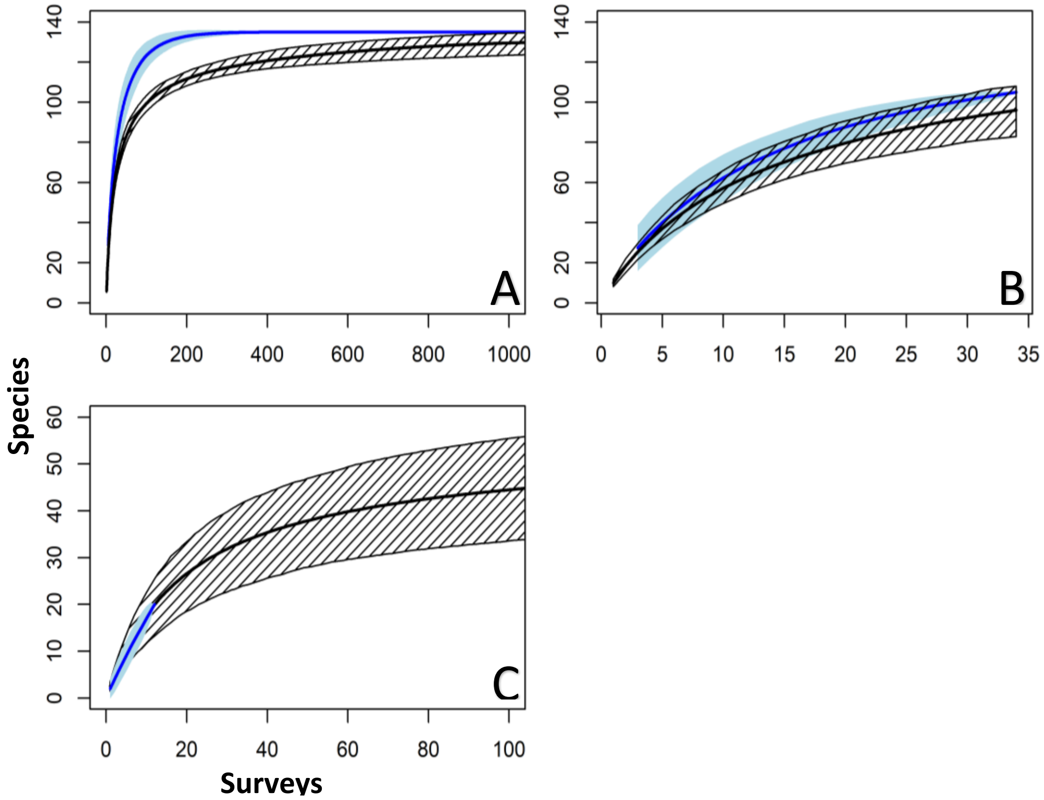

Figure 4-4: Observed (blue) and modelled (black) species accumulation (±2SD) with increasing numbers of surveyed stations, (A) within the training data, (B) set aside for validation, and (C) based on an increasing number of replicate surveys conducted at a single survey station.

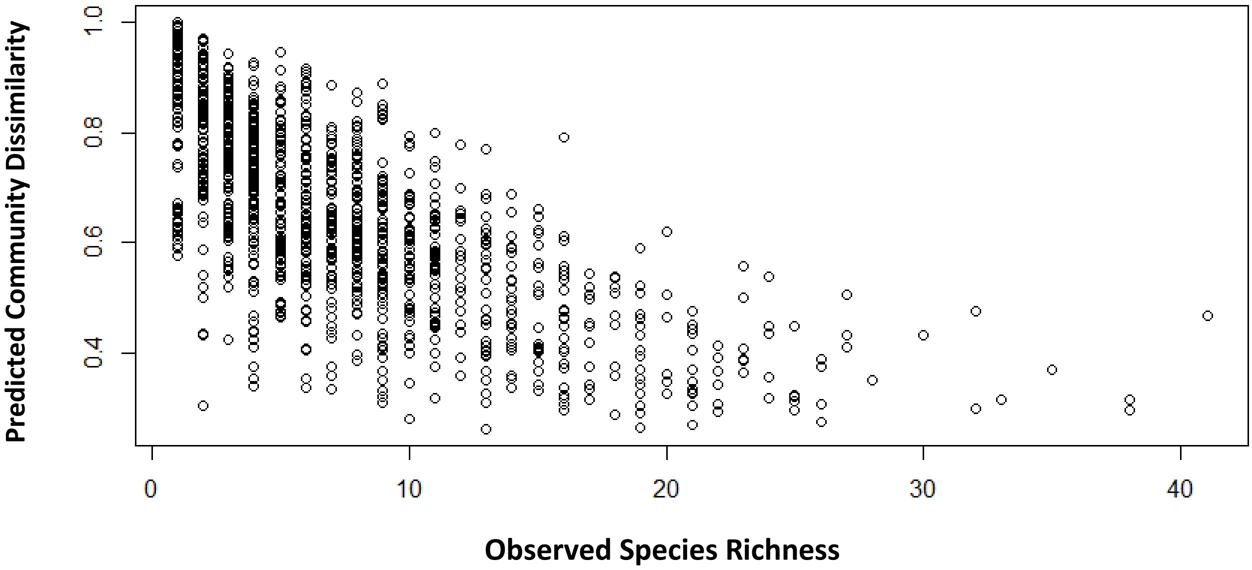

Figure 4-5: Each point represents a survey station. For each station, modelled dissimilarity is the mean dissimilarity between each of the communities produced by the model's 500 posterior samples, and the observed community.

4.4.3 Power to detect change

As a result of low detection probabilities and low mean replicate surveys, it was ascertained that there is a mean probability across all species of 0.20 (±0.24SD) of detecting a 50% reduction in occupancy in a subsequent survey. There is a 70% or greater probability of detecting a 50% reduction in only five of the 135 species modelled. Thus, if the desire is to determine where occupancies are changing by surveying the same area in the future, the current data cannot provide robust evidence.

Figure 4-6 illustrates how different numbers of survey stations and replicates per station affect the median power to observe a reduction in bird occupancy. Each panel represents a different proportional change that is sought to be detected, and each y-axis represents the probability of detecting that change in at least half the bird species. The x-axis is the total number of individual surveys that a study is able to undertake (i.e. one replicate at one station is one survey). Each line represents a different number of sites that the total number of surveys are divided among. For every probability of detecting the panel's proportional change (y-axis), the farthest left line represents the fewest total surveys required, and the most efficient number of sites to survey. For example, if our aims were to have a 70% confidence (0.7 on the y-axis) that we could detect when the occupancy of at least half the bird species had declined by 50% (panel C), the chart suggests that the fewest surveys required would be ~10,000 and spread between 100 to 500 sites. Calculations suggest that the requirements would be three replicate surveys at nearly 50,000 stations, or, if all stations were surveyed 27 times, only 447 stations would be required.

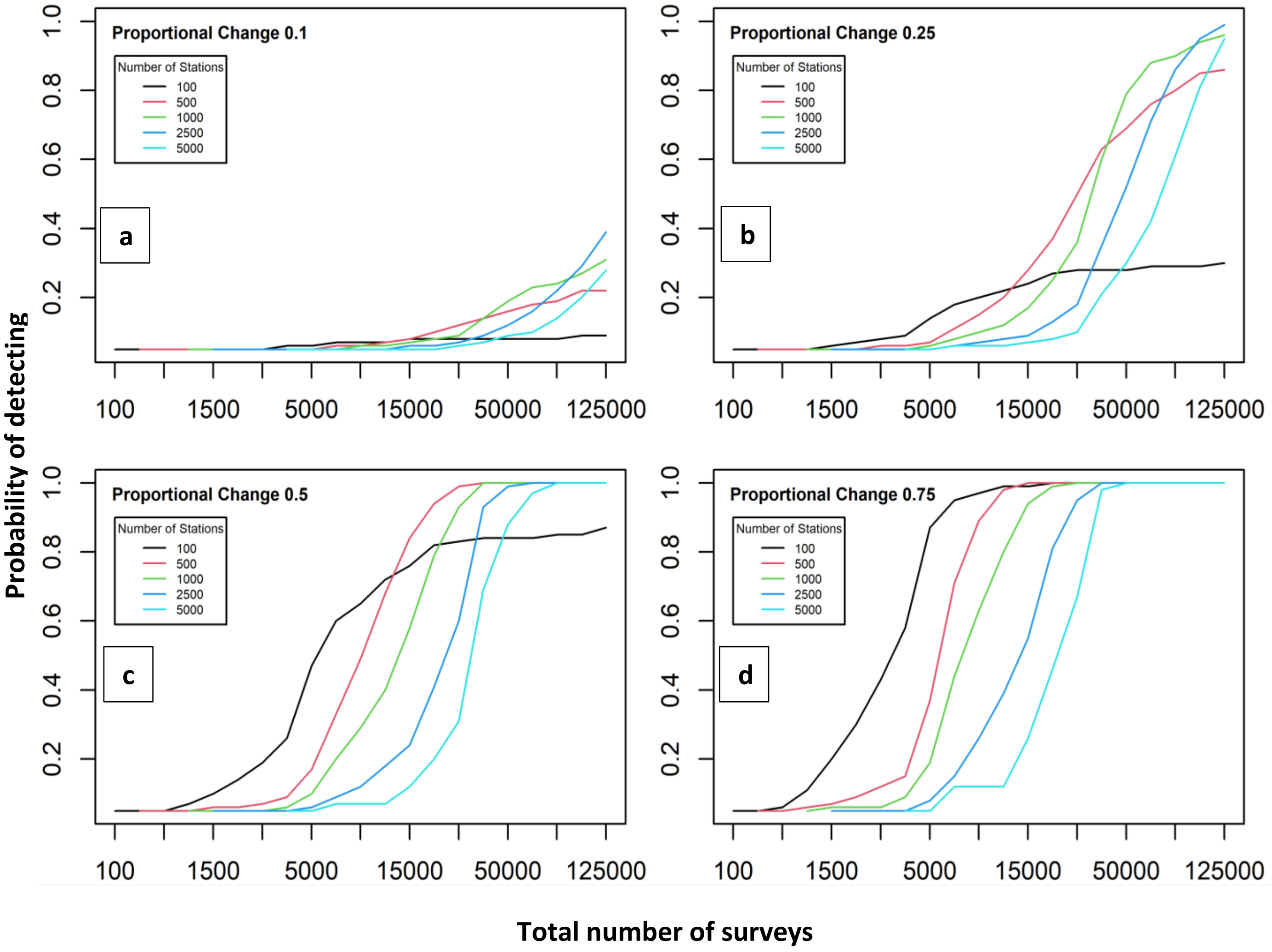

Figure 4-6: The influence of survey effort on the probability of detecting a reduction in occupancy in future surveys. Survey effort was calculated as the number of stations surveyed * number of replicate surveys per station. Each line represents a different number of stations and was calculated as the mean value across all 135 species modelled. Each chart represents a different proportional reduction in occupancy.

4.5 Discussion

Using data from over 3,000 individual surveys conducted in south-eastern Peru, our study suggests that remotely sensed satellite data could extract relevant information about the surrounding habitat to marginally improve on the performance of equivalent landcover models in predicting bird community composition. We also quantified the low detectability of many bird species, highlighting the challenges conservationists face when trying to monitor their rich diversity in rainforests. While the existing survey methodology provided valuable insights into bird communities, its design was not optimised for detecting change over time. Our findings can therefore help inform the development of more efficient survey strategies, leading to enhanced species detection rates. Higher detection rates would allow for a more complete description of each site’s community, improving the power to detect change, and strengthening the performance of RS models in predicting bird communities across the landscape.

4.5.1 Land class and surface reflectance

There are mixed and conflicting results on the effects of forest and farmland on the comparative richness of bird species found in each (Dvořáková et al., 2023; Tu et al., 2020). Our findings, that more bird species were likely to occur in both floodplain and terra firma forests than in agricultural land, may be supported by evidence that suggests that species richness can increase at forest boundaries, particularly where edges are hard, such as where forest and agricultural land meet (Terraube et al., 2016; Willmer et al., 2022). Furthermore, the greater vertical stratification and canopy complexity often found in primary forests are often found to sustain a greater richness, particularly of specialist, range restricted, and endemic bird species compared to secondary forests (Şekercioğlu et al., 2012; Stouffer et al., 2011; Stratford and Stouffer, 2015). However, our analysis suggested that of the species included in the model, secondary forest had a greater species richness than primary forest. This is consistent with findings that where secondary forests have connectivity to primary forests, generalist and disturbance tolerant bird species as well as species that thrive in more open canopy environments are also attracted, thereby increasing species richness (Barlow et al., 2006). Despite the greater species richness frequently found in primary forests, individual surveys may detect more species in secondary forests. This is as a result of the more open habitat found in secondary forests leading to higher bird densities, more vocalisation and improved visibility (Barlow et al., 2007; Peh et al., 2006; Sekercioglu, 2012). The low mean number of replicates conducted in this study could therefore make secondary forests appear richer in species than primary forests, simply because of the improved visibility to observers.

Secondary forests require the initial loss of primary forests, and localised deforestation can increase the richness of bird species. When relatively small areas of primary forest are destroyed, the occupants of that area move initially to the adjacent forest, thereby temporarily increasing its richness (Rutt et al., 2019). Additionally, as habitat heterogeneity increases with slightly higher levels of patchwork deforestation, generalist species, and those adapted to living in secondary forests quickly colonise the deforested area (Rutt et al., 2019). Our findings are consistent with these observations and suggest that deforestation within a 5km radius had a positive effect on mean occupancy. This may be attributed to broader habitat heterogeneity, which at spatial scales of several kilometres, influences community composition as species dispersal, habitat connectivity, and edge effects become more pronounced at this range (Laurance et al., 2011; Thornton et al., 2011). That our analysis did not identify significant support for an influence of deforestation at distances under 5 km could suggest that not enough habitat heterogeneity is created at that level to significantly impact species richness. Furthermore, we have classed deforestation as a lack of primary forest rather than lack of any forest. Any deforested area may not be kept clear, but may have regrown into secondary forest, and as previously discussed, secondary forest was also shown to hold higher species richness. Classification inconsistencies may also exist, particularly where deforestation occurred before 2000. As a result, some primary forests included in this study may have undergone historical degradation. Further classification inconsistencies may also result from the omission of water bodies from analysis. Within floodplain forests, water levels can fluctuate seasonally (Hess et al., 2015) and edge effects may also occur at interfaces between large water bodies and forests (Laurance et al., 2001). These classification limitations should also be considered when interpreting differences in species occupancy across forest types. Since surveys were conducted across multiple seasons and years, bird detectability may have varied due to seasonal fluctuations or localised extinctions and introductions (Blake and Loiselle, 2015; Stouffer et al., 2011). Moreover, seasonal differences are most pronounced in forest fragments and secondary forest, habitats that are less buffered from the changing seasons (Rutt and Stouffer, 2021). Additionally, surveys were performed by multiple observers with varying levels of experience, and observer skill can significantly influence bird species detectability (Robinson and Curtis, 2020). Consequently, some of the variation in detection probabilities may be attributed to differences in surveyor rather than actual ecological differences.

Diversity in surface reflectance values has also been used to predict tropical forest diversity (Peña-Lara et al., 2022) and bird species richness found within them (Suttidate et al., 2019). Our results reflect similar findings, and are in line with studies showing that RS data performed as well as, if not better than, land class in predicting bird communities (Regos et al., 2020; Sheeren et al., 2014). The season of image measurement and resolution of imagery were also found to be important for modelling bird communities in temperate regions (Regos et al., 2020; Sheeren et al., 2014). The seasonal and cumulative productivity of tropical forests ascertained through satellite imagery were also found to be associated with bird species richness (Suttidate et al., 2019). However, in our study area, excessive cloud cover precluded the use of seasonality, as images were reduced to annual composites, and mean reflectance values are calculated over thousands of square metres, making finer resolution imagery unlikely to improve results.

Despite these challenges, our study has shown that RS data have the capacity to improve on land class in predicting bird communities in tropical forests. When the first two PCAs were plotted against each other, the land classes were not tightly grouped and had much overlap, yet two broad groups could be differentiated: secondary forest/agriculture and primary forest/terra-firma/floodplain. Within those two groups, secondary forest had a large overlap with agricultural land. As secondary forest was defined by its status as land that had been deforested since 2000, this could be an indication that, in this area of the Peruvian Amazon, deforested land primarily becomes agricultural land, rather than regenerating into secondary forest (Cruz et al., 2023; Marcus et al., 2020). In the other group, primary forest had a far greater overlap with terra-firma and floodplain than with agricultural land. However, there was still a sizeable overlap between all groups. One possible explanation is that there are discrepancies brought about by course scale, out of date information or human error in habitat classification (Gallardo-Cruz et al., 2024; Morrison, 2016; Ploton et al., 2020). There is also the possibility that the classes fail to capture, or oversimplify the finer scale variation seen by RS data. In fact, secondary descriptors of structural variation at the local level, such as deadwood volume and tree girth, can also influence reflectance values, and when included in conjunction with broad habitat type, have been shown to define species richness far better than habitat type alone (Basile et al., 2021). This highlights a key advantage of using continuous RS data, which can describe gradual changes in vegetation structure better than the rigid structure of predefined classes. The challenge for future research will be in deriving RS variables that can act as more ecologically meaningful proxies for environmental variation that can more readily differentiate between less obvious variations in surface structure.

4.5.2 Realities and practicalities of observing change

Whether models use habitat or RS variables to estimate ecological communities over large areas, their success is limited by the quality and quantity of the data with which they are trained (Aubry et al., 2017; McCord et al., 2021). Our analysis indicates that detection probabilities were so low that substantial effort may be required to detect the majority of species present, raising concerns that our data were insufficient. However, of the approximately 760 bird species thought to occur in the Tambopata area (Lepage, 2024), our study recorded nearly half. This, in terms of both number and proportion of species identified, aligns with findings from other surveys in tropical forests (Alvarez-Berríos et al., 2016; Anderson, 2009; Martin et al., 2017). Due to the low detection rates, individual surveys inherently underestimated community composition, and communities across the landscape were likely to be more similar than observations from a limited number of surveys would suggest. Nevertheless, the species modelled exhibited patterns of grouping based on both forest habitat descriptors and RS variables, suggesting that they represented more than just random observations. This indicates that, despite describing approximately a half of the full community, our results were reliable.

As sampling efficiency or effort increased, a greater portion of the local community was observed, and the similarities between predicted and observed communities increased. Thus, improving detection rates is essential to increasing the proportion of the community observed and improving the predictive output of models. Increasing cumulative detection rates to 80% has been shown to negate the importance of detection on model performance (Steenweg et al., 2019). Yet we have shown that it would require a huge effort to reliably observe changes in occupancy for many species. Thus, efficient and accurate surveys are critical to provide high quality data on which to train RS models, especially as ongoing surveys are needed to monitor the effects of conservation efforts on habitat change over space or time. The type of survey conducted affects the raw detection rate, and we agree with

Mulvaney & Cherry, (2020) that point counts increase mean detection over net counts. However, these findings are not universal, as it has been shown that within tropical forests, the most effective survey method varies with ecosystem, such as whether the study is within lowland or cloud forest (Martin et al., 2017). Similarly, transect counts and canopy counts add further depth to the proportions of small-bodied species, canopy and mid-storey species that are detected (Anderson, 2009; Robinson and Curtis, 2020). Due to the relatively low community overlap observed between methods found in our study, and aligning with other studies, we would suggest that the use of multiple methods is necessary to improve overall diversity assessment (Anderson, 2009; Martin et al., 2017; Robinson and Curtis, 2020). Further to the survey methodology lies the challenge of recruiting appropriately trained surveyors who are familiar with the bird species of the area, which is a necessity in conducting effective surveys in complex and rich ecosystems such as tropical forests (Robinson and Curtis, 2020). Our results suggest that to effectively detect changes in community composition, it is necessary to perform far more replicate surveys than were performed (Robinson and Curtis, 2020).

4.5.3 Future potential

Forest degradation in Peru has been shown to increase the amount of carbon released into the atmosphere by almost 50% (Asner et al., 2010). Understandably then, protecting forests and their carbon sinks to mitigate climate change has been a focus of tropical conservation, and government policies and voluntary carbon offset conservation efforts such as REDD+ have successfully slowed the rate of deforestation in recent years and have played a role in mitigating climate change (Smith et al., 2014). To ensure that the implementation of carbon conservation efforts, such as REDD+, also have a positive impact on biodiversity, ecological monitoring of forests should be aligned with the conservation objectives of the area (Harrison et al., 2012). While REDD+ has had some success in protecting forests globally, the level of success in the Amazon may be overstated due to the overestimation of baseline deforestation (West et al., 2020), highlighting the need to validate the costs of voluntary carbon offset credits. The impacts of degradation are not fixed and can be improved by restoration or worsened by repetition and should be measured at a landscape rather than patch scale (Betts et al., 2024). Measuring community diversity at landscape scales may require impractical levels of monitoring, but patch scale changes in habitat should in principle be observable by Earth observation. Should sufficient, consistent, reliably representative biodiversity data be available to calibrate models fitted with RS data, biodiversity responses to changing land use can be identified. Policy and management should judge the results of practices at the point of implementation rather than judge the practices conceptually as a whole, and by mapping species diversity in conjunction with other measures of landscape health, trends in forest degradation and regeneration over time can be identified (Betts et al., 2024).

Despite the challenges posed by generally low detection rates, our study demonstrated that the occupancy of a third of the bird species with sufficient detections in the Peruvian Amazon can be predicted with reasonable confidence. However, these low detection rates also limit the ability of models to detect changes within these communities over time. While RS models may not yet be sufficiently robust for ongoing monitoring, reporting and verification, there is no evidence that field-based methods alone will be able to do this at the scale required. Survey data can help identify spatial variations in community composition, but detecting temporal changes would require detailed count data at a level that may be impractical (Stephenson, 2020). Our power analysis revealed that with the current monitoring structure, even if a species population were to half, we would fail to detect this four out of five times. The suggestion that declines of many species may be substantial before being reliably detected has major implications for conservation monitoring using existing survey design. Improving detection rates will increase the power of future surveys to detect change. Given that resources are limited, future surveys should focus on fewer locations, prioritised across an ecological gradient of interest, with a higher number of replicates; results suggest at least 20 per site. This is particularly critical as even basic inventories of avian diversity across much of the tropics are lacking (Robinson and Curtis, 2020). Establishing baseline datasets is essential for enabling future repeat surveys, which could provide valuable insights into how and where tropical bird communities are changing (Magurran et al., 2010). Building on the findings of this study by refining survey design and aligning research with specific conservation objectives will be key to improving biodiversity monitoring in the region so that decisions are based on detectable trends rather than snapshots of biodiversity (Nichols and Williams, 2006).

Code and data used in the analysis of this chapter are available at

https://github.com/AndrewCSlater/Chapter---Peru

5. Linking fire, logging and bird communities with Earth observation in the Brazilian Amazon.

5.1 Abstract

Almost half of the natural forests within the Amazon basin have been cleared or degraded by fire or selective logging. The negative impact of forest disturbance on biodiversity is well documented; however, there is a paucity of knowledge about the spatial and temporal variation in biodiversity in response to the type and level of disturbance. While Earth observation (EO) satellites can be used to infer land use and elements of forest structure with consistency and repeatability, their ability to predict associated biodiversity remains uncertain. This study used biodiversity data gathered by passive acoustic monitoring, to investigate how bird communities vary with forest disturbance, and tested whether EO derived metrics could predict bird occupancy and community composition.

Bird communities identified from sound samples were described at 29 sites, across five forest gradients within the Sustainable Amazon Network, in the state of Pará in northern Brazil. Modelling the 163 most frequently observed species, it was found that species richness was significantly lower in secondary forest compared to undisturbed primary forest. Predictive models fitted with EO-derived metrics achieved only moderate success, with a mean area under the curve (AUC) values between 0.56 and 0.61 for bird occupancy, with the number of individual species exhibiting high predictive ability (AUC>0.7) ranging between 23 and 33. The Bray-Curtis dissimilarity values for community composition was ~0.45 for all models.

No single model was consistently superior, but the most effective model was fit with proportions of forest lost and burnt derived from pre-processed Earth observation data. However, the model fit with categorical forest gradient exhibited a near equivalent performance. Models using radar derived variables showed potential utility, while models fitted with directly observed reflectance variables were slightly less informative, as the data were significantly influenced by cloud cover. Acoustic monitoring yielded high detection rates and described biodiversity equally well across all sites. However, the findings of this study

suggest that while EO data an differentiate communities equivalently to forest disturbance categories, its ability to model finer-scale variations in bird communities is currently limited. Expanding the number of survey sites could improve model accuracy. With sufficient training data, EO models may offer a scalable tool for biodiversity monitoring and conservation planning.

5.2 Introduction

Tropical forests account for approximately half of terrestrial carbon uptake and up to two-thirds of the carbon sink in biomass, soil, and deadwood of all forests worldwide (Hubau et al., 2020; Pan et al., 2011), and are the most biologically diverse areas on Earth (Bradshaw et al., 2009; Gibson et al., 2011). Nevertheless, the practices of clear-cut logging and commercial agriculture in conjunction with both anthropogenic and natural fires have resulted in the alteration or complete destruction of forest habitats. This is evidenced by the fact that over 17% of the Amazon Basin, which is the world’s largest continuous tropical forest, has been deforested, and that 38% of the remaining area has been lost or degraded (Lapola et al., 2023; Potapov et al., 2017; van Wees et al., 2021; Vergara et al., 2022). The disturbance and deforestation of tropical forests has a negative impact on biodiversity (Barlow et al., 2016). The temporality of disturbances, such as repeated fires or more recent logging, can exacerbate the negative effect on carbon sequestration and on biodiversity (Burivalova et al., 2015; Rappaport et al., 2022, 2018).

Although it is well established that forest disturbance influences biodiversity, the extent to which species composition and ecosystem function vary spatially and temporally remains poorly understood. One of the primary challenges to biodiversity monitoring is the significant time and expense required to comprehensively survey large areas (Rocchini et al., 2016; Skidmore et al., 2015). However, remote sensing encompasses a growing set of techniques that enable automated broad-scale monitoring of environments with limited expense. Earth observation (EO) by satellites is considered an important approach to vegetation monitoring (Skidmore et al., 2015). While satellites may not be able to directly observe biodiversity, they can be used to provide proxies of forest structure and to measure deforestation (Clark et al., 2021). Similarly, passive acoustic monitoring is becoming increasingly prevalent in monitoring soundscapes and the inference of the biodiversity of vocal faunal communities including birds, insects and anurans (Metcalf et al., 2022). The integration of fine-scale biodiversity surveys with EO descriptors of the Earth's surface in statistical models, enables the analysis of community composition and turnover of target taxa (Leitão et al., 2015). EO derived forest structure metrics such as vegetation indices and canopy cover can then serve as indirect indicators of habitat quality and faunal communities (Turner et al., 2003). However, while EO can detect some structural aspects of forests, it does not directly measure the faunal biodiversity found within them, and its effectiveness of EO in capturing fine scale variations in biodiversity varies on habitat complexity and species detectability (Jetz et al., 2016).

Bird communities are closely linked to forest habitat structure and are sensitive to changes caused by deforestation and fire, making them effective bioindicators to monitor forest degradation change which is detectable by EO (Barlow et al., 2016; Ferraz et al., 2007). Importantly, species occupancy and community composition respond to environmental changes at different spatial scales, meaning that EO monitoring must consider habitat features at local and broader scales (Lausch et al., 2015). If EO information is demonstrated to be a reliable predictor of bird communities, it could be employed to enhance the ongoing monitoring of ecosystem condition across the region. This is particularly relevant, as El Niño events and their associated forest fires are anticipated to increase in frequency (Burton et al., 2020; Wang et al., 2017). While deforestation rates rose between 2018 and 2021 (Silva Junior et al., 2020), recent data indicate a decline since 2022 (WWF-Brazil, 2024). Despite these fluctuations, deforestation remains a key driver of structural change and thus the biodiversity of the wider ecosystem (Flores et al., 2024), reinforcing the need for scalable monitoring.

Birds represent one of the most extensively studied taxa within the Amazonian forests (Rappaport et al., 2022). They have been employed as indicators to assess the impacts of environmental change, including the evaluation of reforestation techniques (Barros et al., 2022), and the impact of patch size reduction within the Brazilian Amazon (Bregman et al., 2015). This study employs data on bird communities collected via passive acoustic monitoring to investigate whether such communities differ across forest gradients that are determined by whether a survey site has been logged and/or burnt. Furthermore, the study assesses the capacity of forest gradients to predict the bird communities that inhabit them. Subsequently, the study evaluates whether EO derived forest structure metrics can effectively predict the same bird communities. We hypothesise that bird community composition will differ significantly between sites that have been logged, burnt and undisturbed, and that EO metrics will be significant predictors of those differences. If EO derived metrics can reliably predict bird communities, they could provide a scalable tool for biodiversity monitoring, aiding conservation planning in tropical forests that are affected by fire and logging.

5.3 Materials and Methods

5.3.1 Study Area

The data used in this study were gathered across permanent terra-firma transects of the Sustainable Amazon Network (Gardner et al., 2013), situated to the south of the city of Santarém in the state of Pará in the eastern Brazilian Amazon (Figure 5-1). A total of 29 sites were distributed across five forest classes. One class was unburnt secondary forest that had grown since the clear-cutting of primary forest (n=3). Four classes were identified within primary forest: undisturbed (n=5), logged but unburnt (n=4), burnt but not logged (n=5), and logged and burnt (n=12).

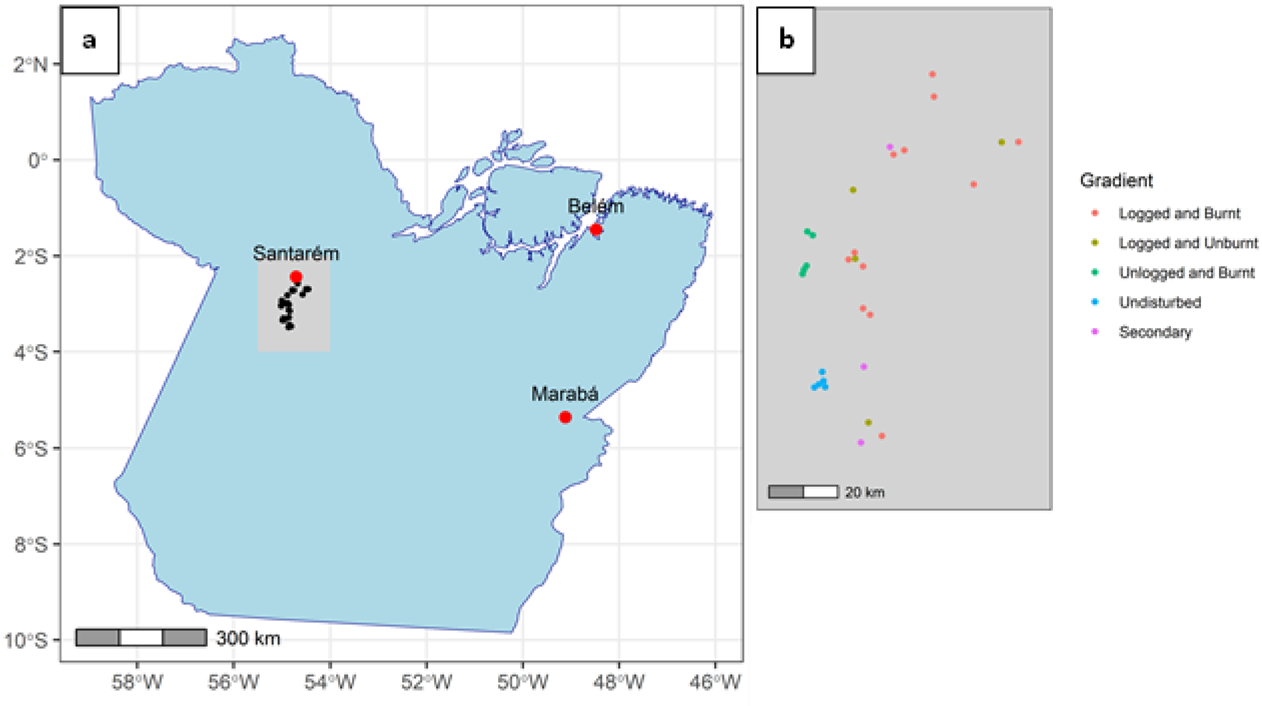

Figure 5-1: (a) The geographical distribution of the survey sites in the state of Pará, represented as black dots within a grey box. Major cities are indicated by red spots. The inset shows the state of Pará within Brazil. (b) The distribution and forest gradient of the survey sites as illustrated in the grey box in panel (a).

5.3.2 Species Data

All surveys were conducted over a 64-day period, with varying numbers of days per site, ranging from a minimum of six to a maximum of 22 days, with an average of 16 days (standard deviation (SD) = 3.37 days). Sites were monitored by audio recordings 24-hours per day. For each site, a total of one hour of recording was then sub-sampled by randomly extracting 240, non-overlapping, 15-second clips from a 2.5-hour window that began 30 minutes before dawn. The audio clips were analysed by a specialist to identify bird species heard. The timing aligns with peak avian vocal activity and can be considered an automated replication of traditional point counts. Consequently, species that are less vocal or active outside of the sampled window may be underrepresented (Martin et al., 2017). Although all sites were subsampled by an equal amount of time, some clips were subsequently removed due to the presence of rainfall noise which rendered the data unusable. The sites were surveyed for a minimum of 2880 seconds, a maximum of 3600 seconds, and a mean of 3423 seconds (SD=188) (Metcalf et al., 2022, 2021). All 15-second clips recorded within a calendar day were treated as a single replicate survey. It was anticipated that the variation in the effort per replicate (number of seconds sampled per day) would influence the number of species recorded. Consequently, survey time was included in the analysis as a detection variable.

A total of 224 species were identified. The mean number of sites in which species were found was 8.7 (SD=7.7). The minimum number of species recorded at a site was 34, with a maximum of 91 and a mean of 67.3 (SD=14.8). A total of 61 species were recorded at fewer than three sites and were therefore excluded from further investigation, which left 163 species for analysis.

5.3.3 Remotely sensed data

The presence of cloud cover over tropical forests can make the acquisition of clear, temporally relevant, images challenging (Quiñones et al., 2007). Planet satellites provide a daily repeat and provide a pre-processed monthly base map of surface reflectance, which masks cloud and is normalised for analysis. The most suitable image available that coincided with the survey period was the monthly image from June 2018, and thus this was utilised. However, despite this, the image still exhibited visible signs of atmospheric haze. For the purposes of this study, composite or cloud-masked images from Landsat and Sentinel-2 were found to be lacking in data or to exhibit significant distortion due to cloud cover, rendering them unsuitable for the intended analysis.

In addition to individual pixel values, the spatial patterns in pixel properties are also expected to convey information about forest structure (Rocchini et al., 2004). Grey Level Co-occurrence Matrices (GLCM) describe the pattern of co-occurring neighbouring pixel values within a moving window across a single raster layer (Zhou et al., 2017). Due to the presence of a slight haze across the Planet image, we investigated the potential for patterns of pixel values to provide a more accurate representation of the forest canopy than raw pixel values. In order to achieve this, we calculated GLCMs for each of the four bands of the Planet image. A 9x9 pixel window (~45m) was used, which upon visual inspection of satellite images, appeared to effectively cover a uniform forest type with minimal influence by neighbouring types on the window (Hall-Beyer, 2017). A variety of methods exist for describing pixel patterns, and we calculated the contrast, entropy, and mean values (Hall-Beyer, 2017; Haralick et al., 1973). The GLCM calculations were performed in R using the “glcm” package (Zvoleff, 2020).

In contrast to optical earth observation, Sentinel-1 are synthetic-aperture radar satellites that actively emit polarised pulses and measure the strength and polarity of the returned signals. The influence of canopy cover, structure and water content on radar reflectance is well documented (Ustin and Jacquemoud, 2020). In contrast to passive reflectance, radar imagery has the capacity to marginally penetrate the forest canopy to describe the structure beneath, and importantly, is unaffected by cloud cover. Studies have demonstrated the utility of radar imagery in differentiating between some forest types with differing structural properties but with similar colour reflectance (Numbisi et al., 2019; Solberg et al., 2014), which in turn influence habitat suitability for different bird species. A single, pre-processed mosaic image from 2018 that provided vertical-vertical (VV) and vertical-horizontal (VH) bands at a 10m^2 pixel resolution was downloaded from Google Earth Engine. In addition to EO data temporally linked to the time of surveys, areas of deforestation calculated from historic Landsat data (Hansen et al., 2013), and areas of forest loss due to fire (Tyukavina et al., 2022) were also used. To assess the influence of scale of measurement on model performance, the mean and standard deviation of all pixels within 30m, 250m, 1,000m and 2,500m of each survey site were calculated for all data except for GLCMs, which already summarise the area around the survey site. This approach reflects the known importance of spatial scale and landscape structure on biodiversity, and the scales align with typical resolutions used in RS studies, reflecting local to landscape habitat levels (Lausch et al., 2016; Rocchini et al., 2010).

5.3.4 Statistical Analysis

To investigate the potential for forest type to be able to describe and predict bird communities, a model was fit using the forest gradient as described by the site survey. To explore the ability of RS data in describing and predicting the composition of bird communities, and to compare their abilities with those of the forest gradient, a series of models were constructed using a range of combinations of optical, radar and proportional areas of deforestation at each measurement radius. As some variables within groups were likely to be redundant, and to reduce the probability of overfitting models, and to ensure each model was populated with the same number of covariates, principal component analysis was used to reduce all combinations of variables to two components. The first to components accounted for 92.7% of total variance and fits with the guideline of having at least 10 observations per predictor variable (Vittinghoff and McCulloch, 2007). Furthermore, a model was constructed without occurrence variables to serve as a baseline intercept model for comparison.

To quantify the probability of species detection, and to identify whether the composition of bird communities can be predicted as a function of habitat or RS descriptors, we used spatial factor multi-species occupancy models. The detection of a species in a given survey is dependent on two probabilities. The physical incidence of a species (1 = present, 0 = absent) is derived from its probability of occupancy at a given location, and then, given occurrence, its probability of detection. Occupancy models account for imperfect detection in calculating the probability of occupancy of a species at a survey station. They also provide probabilities of occupancy at all stations, whether observed or not. Furthermore, occupancy models provide detection probabilities for all species at all stations as well as the influence of survey-specific covariates on these probabilities.

A series of models were constructed using the recently released Bayesian occupancy model package, “spOccupancy” (Doser et al., 2022). The “spOccupancy” package allows for the simultaneous modelling of all species, while accounting for autocorrelation between species. In addition, spatial latent factors, which are akin to unmeasured site-specific covariates that decay with distance, can be incorporated to account for residual species correlations. In order to achieve the greatest efficiency and to avoid overfitting, particularly in the case of a limited number of replicate surveys or sites, it is preferable to include fewer factors rather than more. It is not possible to define what spatial factors represent or are driven by, meaning that they provide little benefit to explanatory models. However, when predicting across unsampled areas, spatial factors help account for unmeasured, spatially correlated environmental heterogeneity, thereby enhancing predictive capabilities, and their incorporation when prediction is a primary objective is advised (Doser et al., 2023). Each spatial factor provides a rate of decay value (phi), and 3/phi represents the distance at which stations are no longer autocorrelated for the given factor. All models included two spatial factors in addition to survey time (number of seconds), which was included as a covariate expected to influence detection. Finally, to ascertain the influence of spatial factors on model performance, the most effective RS model was refitted without these factors. Models tested are shown in Table 5-1.

Table 5-1: Models that were tested. All models included survey seconds as a detection covariate. All models, apart from one group of EO models, included two random spatial factors as occurrence covariates. Models fitted with Earth Observation data had all variables reduced to the first two primary components from PCA analysis. Each of the Earth Observation data groups had four models fitted, one each for data measured at 30m, 250m, 1km and 2.5km around the survey points.

	Intercept	Forest Gradient	All EO	EO ex Planet	EO ex Planet & Space	Prop Logged Burnt
No Occurrence Covariates	X					
SURVEY DATA						
Forest Gradient		x				
EARTH OBSERVATION DATA						
GLCM from Planet Imagery (12 variables)			x			
Sentinel-1 Radar Imagery (4 variables)			x	x	x	
Proportion of Deforestation and Burnt area (2 variables)			x	x	x	x
SPATIAL FACTORS						
Two	x	x	x	x		x
DETECTION COVARIATES						
Survey Seconds	x	x	x	x	x	x

Models were fitted with 12,000 iterations, of which the first 2,000 were discarded as burn-in. The remaining iterations were thinned to every tenth one to produce 1,000 posterior samples. Chain convergence is particularly difficult in spatial models, so models were fitted using a single chain, and mixing was assessed visually and using the Geweke diagnostic (Doser, 2023). In fitting the models, it was assumed that there were no false presences, and the study was treated as a single closed season. Each posterior sample provided a probability of occurrence (psi) for each species at each station along with a latent presence/absence (1/0) (z) calculated from this probability. In addition, a probability of detection (given presence) (p) was provided for each species at each station for a single survey. The goodness-of-fit for each model was evaluated by performing a posterior predictive check to calculate a Bayesian p-value, based on the differences between chi-squared values of observed and predicted data per species, grouped across stations. All fitted models had Bayesian p-values of 0.2 indicating an acceptable fit to the data (Hooten and Hobbs, 2015; Kéry and Royle, 2016). Subsequently, models were compared using the widely applicable information criterion (WAIC) (Watanabe, 2010). Despite the well-known differences in performance between explanatory models fit on training data and predictive models fit on withheld datasets, predictive performance on independent data is rarely undertaken with occupancy models (Gould et al., 2019). Our models were validated by predicting communities to sites withheld from training data. To ensure all forest gradients were represented in model validation, five sites were randomly selected to represent one of each forest gradient type. The validation sites withheld from the models represented 17% of all available data. The predictive performance of the models was evaluated at the species level using the area under the receiver operating curve (AUC), and at the site community level by Bray-Curtis's dissimilarity of community composition.

Although models provide a probability of occurrence for every species at every location, when assessing predictive performance against withheld survey data, imperfect detection within those surveys must also be accounted for. For each modelled posterior sample, the probability of detection of a species at a station across all surveys was calculated as $1 - (1 - p)^n$, where p is the probability of detection in a single survey, and n is the number of replicate surveys performed at the station. This was used to generate a binomial detection/non-detection (1 or 0) event for each species at each site. The detection event was multiplied by the predicted probability of occupancy (psi) and latent occupancy (z) of the same species and station. In the event that the latent detection or latent occupancy was equal to zero, it was determined that no detection had occurred. Consequently, the community outputs used for validation were those predicted to have been observed, rather than those predicted to have occurred.

Subsequently, for both observed and modelled communities, a pairwise beta-diversity between survey sites was calculated and a Jaccard dissimilarity matrix was created using the '*betapart::beta.pair*' function. Principal coordinate analysis (PCoA) was conducted on the dissimilarity matrix using the *'vegan::betadisper'* function, and the resulting plot was used to observe the variances by forest gradient. All analyses were conducted using R version 4.1.0 (R Core Team, 2022), on R Studio (R Studio Team 2021).

5.4 Results

There is no consensus regarding the criteria for defining a significant difference in WAIC values, but a difference of two (2), as used with AIC values, is often employed (McElreath, 2020). The WAIC values for all models fell within the range of 36,640 and 37,370, with lower values indicating superior performance. However, the model with the lowest WAIC value (RS data measured at 2.5km) was not the most accurate predictive model. The WAIC range suggests there was only minor variation in model fit and model selection was further guided by predictive accuracy.

As species detectability is not influenced by occurrence covariates or spatial factors, it remains consistent across all models. The median intercept detectability for a single survey across all species was 9.4% (95% CI [7.7%, 11.3%]). The probability of detection was found to be significantly influenced by survey effort, with the mean probability of detection across all surveys increasing to 74.7% (SD=20.5%).

In undisturbed primary forest, the median probability of species occupancy was 49.4% (95% CI [13.6%, 92.7%]). The median occupancy was found to be significantly lower in secondary forest at 22% (95% CI [13.2%, 34.5%]). Additionally, the median probability of occupancy was found to be lower within all gradients of disturbed primary forest, although this was not statistically significant as the 95% CI encompassed both positive and negative values.

5.4.1 Model Validation

Only minor differences were found between all models in their explanatory and predictive powers. The mean AUC of all models when assessed on training data was 0.95, while when predicting to independent validation sites, the AUC ranged from 0.53 to 0.61. The three most effective predictive models were fitted with the following variables: the proportion of forest logged and burnt within 250m of survey sites (mean AUC=0.61), survey-defined forest gradients (mean AUC=0.60), and two PCAs derived from radar data and the proportion of forest logged and burnt within 250m of survey sites (mean AUC=0.59). It is noteworthy that none of the best performing models included data from the optical Planet image. When plotted against each other, the two PCAs highlight a division between sites that were logged or not, but no separation by sites that experienced burning is evident (Figure 5-2). In the context of evaluating the ability to predict individual species, an AUC of >0.7 is considered to be an indicator of good predictive performance (Mandrekar, 2010). Species that achieve this level are referred to as high-AUC species. The three best performing predictive models identified 33, 26 and 30 high-AUC species respectively, out of a total of 154. Despite minimal differences in predictive performance across all models, those that excluded spatial latent factors exhibited the poorest performance in all aspects of validation.

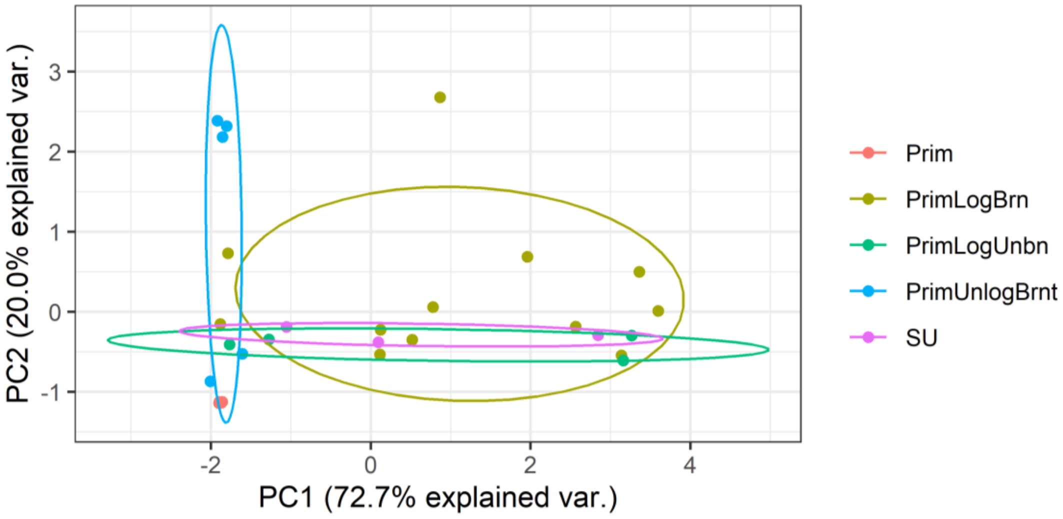

Figure 5-2: Correlation between forest gradient and the PCA variables used in the best RS model. Variables used in creating these PCA were measured within a 250m radius of each survey site and included mean and standard-deviation VV and VH radar bands from Sentinel-1, proportion of forest loss, and proportion of forest burnt. No optical reflectance data were included.

In predicting species richness, the performance of the models was once again comparable. There was a strong overlap between the modelled and observed species accumulation curves. The best RS model demonstrated a minimal degree of under-estimation when applied to training dataset, and a slight degree of over-estimation when predicting to independent sites (Figure 5-3). The models demonstrated a negligible difference in their ability to describe community composition on training data, with a mean Bray-Curtis dissimilarity across sites ranging between 0.158 and 0.167. In the independent validation sites, the mean Bray-Curtis dissimilarities across all models were between 0.445 and 0.478.

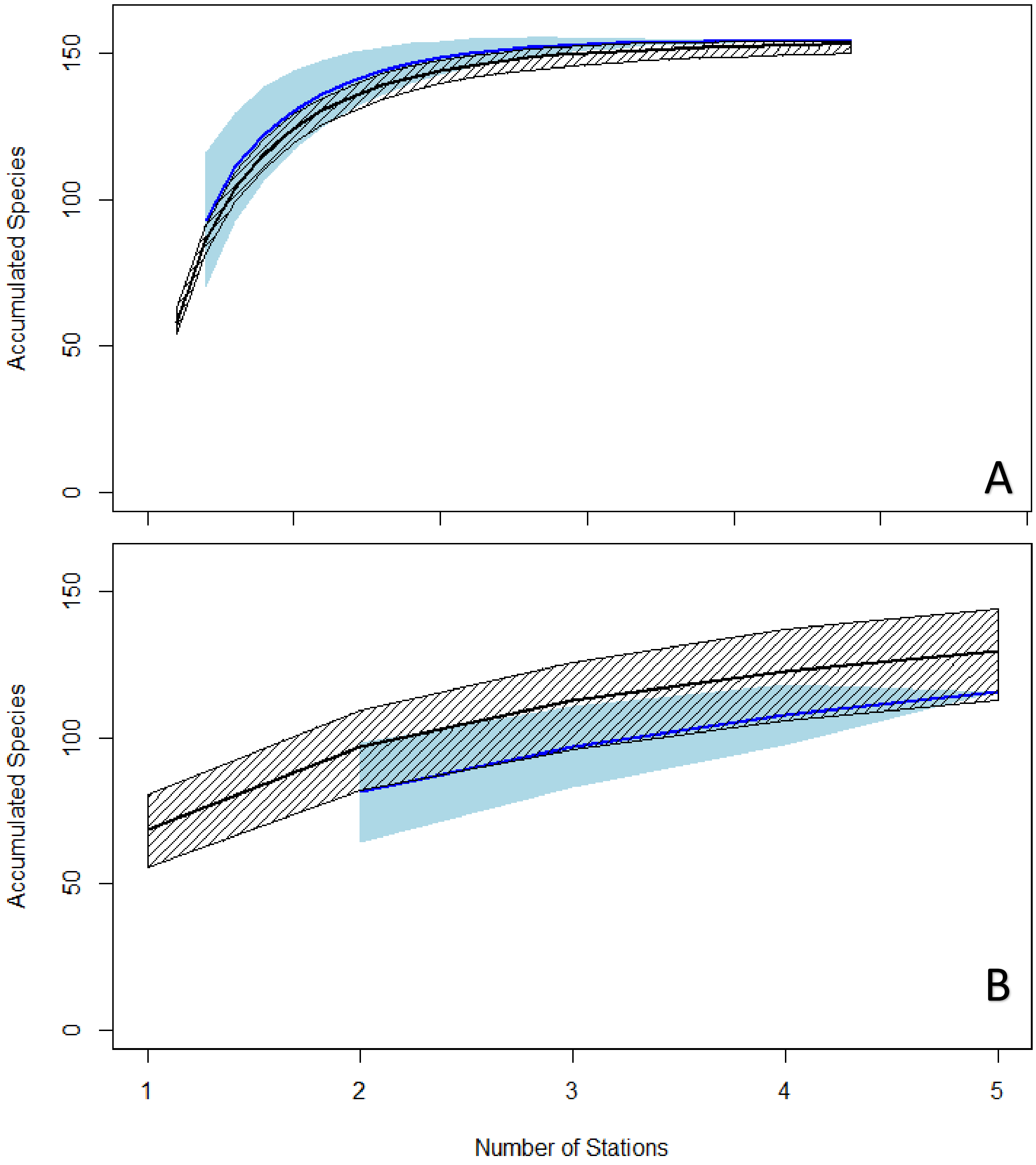

Figure 5-3: Observed (black) and modelled (blue) species accumulation (+/-2SD) by increasing numbers of stations surveyed, (A) within the training data, (B) set aside for validation.

5.4.2 Community richness and composition

A significant disparity between the observed richness of secondary and undisturbed primary forest (p=0.025) was noted when using a Pairwise Wilcoxon's test, although no significant difference was identified between any other gradients. Conversely, a significant difference in predicted species richness was observed between all forest gradients (p<0.001 for each pair) using a Wilcoxon's test. No significant difference was found between the predicted and observed richness's of either the Primary Untouched or the Primary Logged and Burnt gradients (Kruskal-Wallis's tests p=0.30 and 0.20 respectively). However, there were significant differences between the observed and predicted mean richness of the remaining three gradients (p<=0.05) (Figure 5-4).

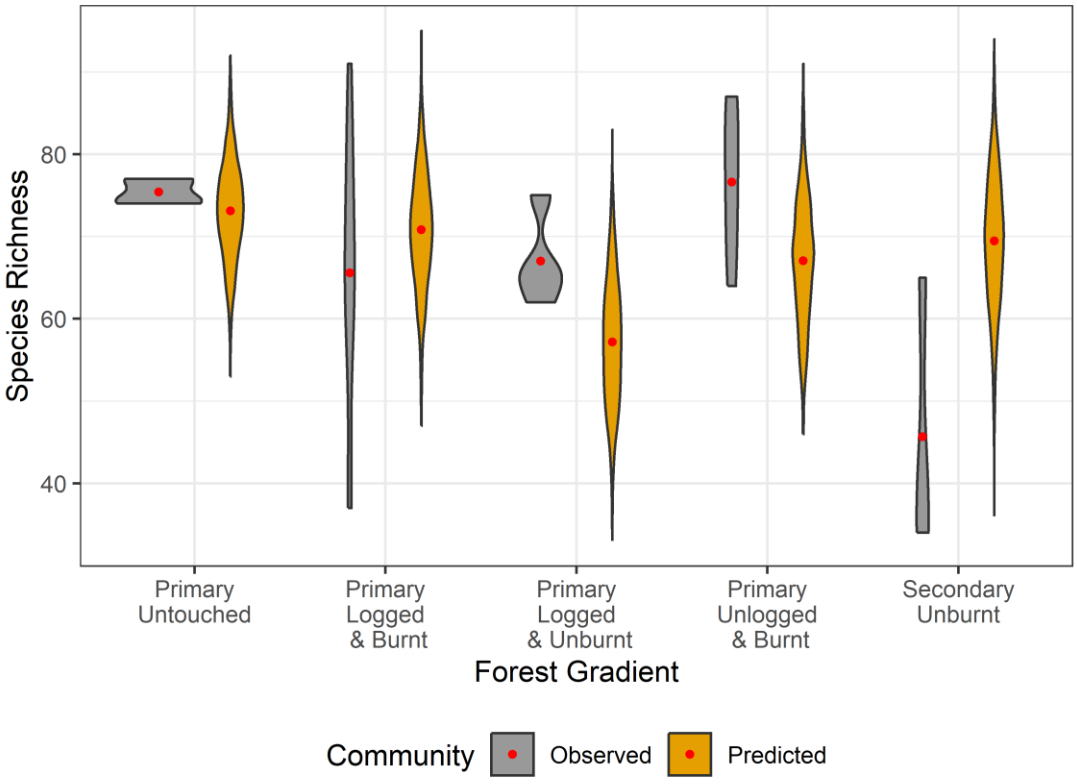

Figure 5-4: Species richness of observed and predicted bird communities split by forest gradient; red dots indicate mean values. Observed values are from all 29 surveys, predicted values are from 1,000 posterior samples of the best model for each of the five validation sites.

Principal coordinate analysis of the pairwise Jaccard dissimilarity matrix of community composition indicates that there is some variation in community composition between forest gradients for both observed and predicted communities (Figure 5-5). A significant difference was identified between the Undisturbed Primary and Logged and the Burnt Primary Forest types using Tukey's Honest Significant Difference (HSD) tests (p=0.033). However, similar tests on predicted communities indicate that significant differences exist between all gradients (p<> 0.019 and 0.000).

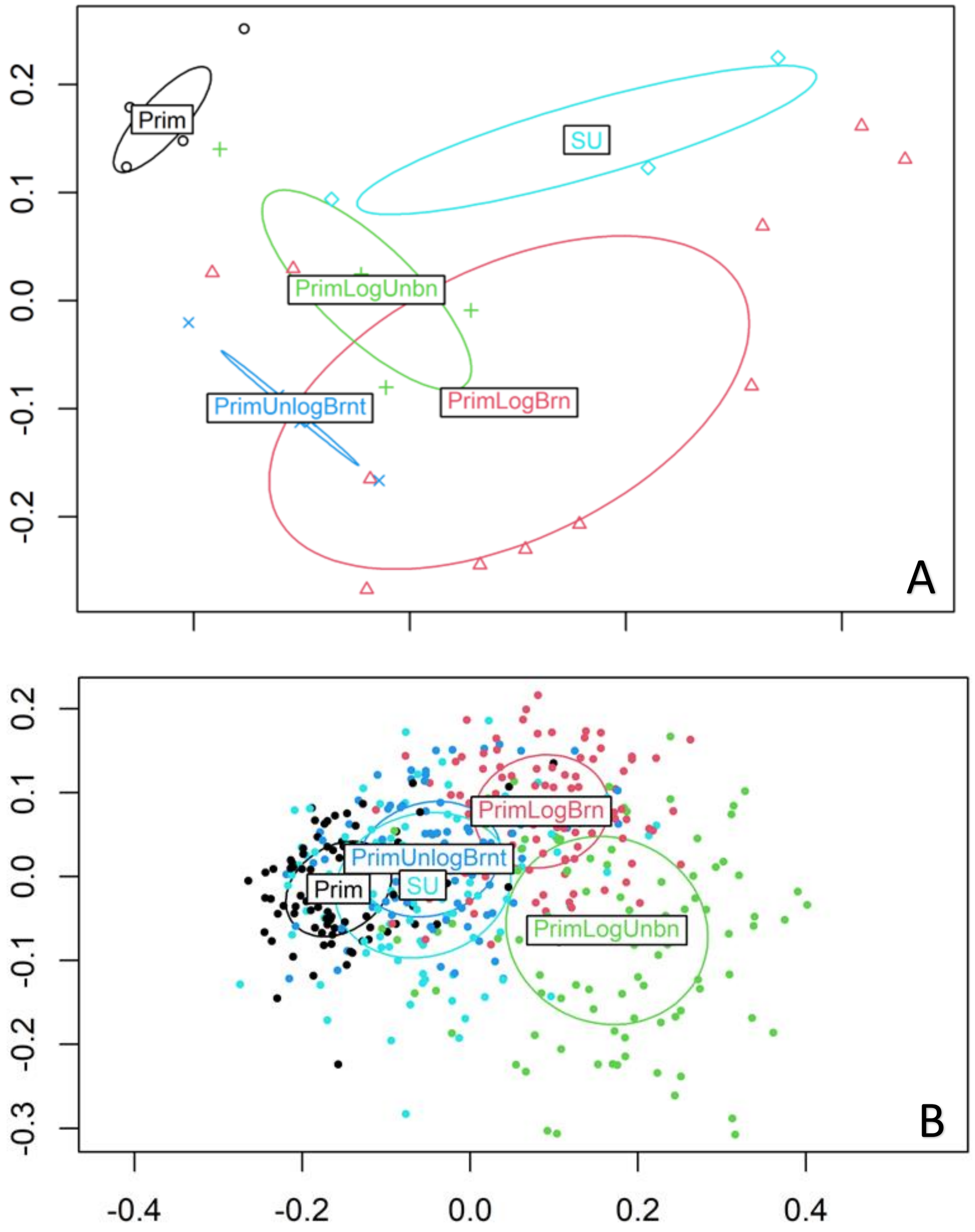

Figure 5-5: The variation in community composition calculated from a PCoA of pairwise Jaccard dissimilarities for (A) the 29 observed communities and (B) 5 validation communities predicted 1,000 times each. Both charts show the forest gradient of each point, and ellipses show one (1) standard deviation around the medians of each forest gradient.

5.5 Discussion

This study analysed bird communities identified by passive acoustic monitoring within forests of the North-Brazilian Amazon, as birds serve as bioindicators of habitat change due to their sensitivity to deforestation and fragmentation (Barlow et al., 2016; Ferraz et al., 2007). The 29 survey points were distributed across five distinct forest gradients, defined by the occurrence of logging and/or burning. In order to ascertain the influence of forest gradient on bird communities, and to evaluate the capacity RS data to predict these communities, joint species occupancy models were used. Models fitted with RS data demonstrated satisfactory predictive performance, but no single model was overwhelmingly superior as there was minimal variation in the predictive capacity across all models. Despite this, inferences regarding the correlation between RS data and forest disturbance, and their influence on the structure of bird communities were possible.

5.5.1 Impact of Forest Disturbance

The loss of primary forests has been demonstrated to result in an increase in carbon gases and cause the loss of carbon sequestration and floral biodiversity. While land classes such as primary and secondary forest may be used to define and estimate the location, rates and levels of deforestation, plantation growth and carbon sequestration (Quiñones et al., 2007), they are insufficient for determining the biodiversity of an ecosystem. There are few studies that link faunal communities with forest gradient, but our findings indicated that the richness of bird species was significantly lower in secondary forest than undisturbed primary forest. This challenges the notion that secondary forests serve as a means of mitigating the negative effects of deforestation (Elias et al., 2020; Esquivel-Muelbert et al., 2019; G. D. Lennox et al., 2018; Matos et al., 2020). It may, therefore, be inferred that secondary forests are less beneficial at maintaining biodiversity than previously thought (Smith et al., 2020). Similarly, when the proportion of forest lost and forest burnt were the only covariates, occupancy was also significantly lower in areas that had experienced higher levels of burning, corresponding with the generally reduced occupancy found in disturbed forest gradients. These findings are consistent with studies that show habitat loss and degradation in tropical forests are primary drivers of occupancy change (Barlow et al., 2016). In contrast, and with the exception of the smallest 30m measurement, mean occupancy was significantly higher with greater proportional forest loss in lands surrounding the survey site. The forest gradients represent point values for the survey sites and, similarly to the 30m forest loss measurement, do not significantly influence occupancy, whereas proportional forest loss, measured at greater radii around each site does. However, in this instance, the greater predictive performance of models incorporating disturbance within 250m rather than at 1km or 2.5km of a survey point suggests that occupancy is more sensitive to fine than coarse scale landscape change. These findings align with previous research showing that species occupancy and community composition respond to environmental changes at different spatial scales (Lausch et al., 2015). Forest fragmentation at broader spatial scales can lead to edge effect species turnover (Pfeifer et al., 2017). Some species exhibit strong site-level associations, while others respond to fragmentation and land-use patterns at the landscape scale (McGarigal et al., 2016). Undisturbed primary forest sites were situated more than two kilometres from other forest gradients, landscape types and forest edges. Conversely, all disturbed sites were situated within patchwork landscapes of differing forest types and agricultural lands and were only 100s of metres from a forest edge. It can be observed that there is a discrepancy between the influence of forest gradient and proportional forest loss on occupancy. This appears to be influenced by the scale of measurement and possibly by neighbouring land classes. There are mixed and conflicting results on the effects of forest type and farmland on the comparative richness of bird species found in each (Dvořáková et al., 2023; Tu et al., 2020). Edge effects may occur at the interface between secondary and primary forests. Studies have demonstrated that species richness can increase at forest boundaries, particularly where edges are abrupt, such as the transition from forest to agricultural land (Terraube et al., 2016; Willmer et al., 2022).

Suggesting that deforestation is beneficial to biodiversity seems counterintuitive, and community composition is a measure of environmental health that goes beyond species richness. Habitat heterogeneity and edge effects have been shown to influence species composition (Haddad et al., 2015; Laurance et al., 2011). Our findings follow this pattern, as although mostly non-significant, variations in community composition between gradients seem evident when plotted, inferring there may be some level of influence by logging and or burning on bird communities. This is further affirmed by acoustic indices, which have been shown to accurately predict the same forest gradients (Metcalf et al., 2021). If acoustic indices are heavily influenced by bird calls, this implies that each forest class must have a discernibly different community composition. However, these are only provisional inferences. In order to confirm the true impact of the effects of logging and burning on bird communities, it would be necessary to monitor whether communities change over time in conjunction with forest disturbance (Magurran et al., 2010).

5.5.2 Correlation between forest gradient and RS data

Although predicting forest gradient was not the goal of the study, a correlation with RS values was anticipated. The results demonstrate that PCA1 is capable of differentiating between sites that have been logged and those that have not. In contrast, PCA2 exhibits a graduated distribution, with the lowest values observed in unburnt sites. The PCA values of the sites within unburnt secondary forest and logged but unburnt primary forest are indistinguishable, indicating that the reflectance values used are similar for both gradients. Secondary forest is defined as a forest that has been clear-cut and has recovered to forest status. In contrast, logging refers to the harvesting of trees within a primary forest without clear-cutting and encompasses a range of proportional tree loss. Differentiating between these forest classes using remote sensing has been challenging (Mitchard et al., 2011). The richness and community composition of bird species found in these two gradients varies significantly, thus underscoring the necessity to identify RS variables that can more effectively differentiate between them. In unlogged forests, the PCA2 values for pristine primary forest sites are tightly grouped, whereas those for burnt primary forest are broadly spread. Forest that have experienced both logging and burning exhibit a considerably broader range of reflectance values than the other groups. It may be anticipated that the extent of logging and the resulting impact on reflectance may vary between sites. Similarly, the burnt forests did not all experience the same type or timescale of fire. Some forests burnt because of El-Nino in 2015 and some burnt due to other causes prior to 2015. The reflectance values observed across the logged and burnt gradient are therefore broad and overlapping (Bourgoin et al., 2018), with an equally broad spread of community composition. While radar showed some ability to differentiate between forest gradients, its sensitivity to moisture content and canopy density may have obscured fine-scale structural differences and contributed to the observed variation (van Emmerik et al., 2017). LiDAR, which can provide three-dimensional forest structure data has been shown to improve habitat classification (Fayad et al., 2016). However, despite its potential, the physical and financial demands of obtaining timely LiDAR data for continuous monitoring of change reduces its viability (Pereira Mendes and Lim, 2024). Given these constraints, integrating longer-wavelength radar such as NISAR with existing with passive acoustic monitoring may offer a more practical approach to long-term biodiversity monitoring. Consequently, defining the communities within this category is challenging, but given sufficient training data, RS linear descriptors are expected to differentiate between forest gradients more effectively than land class alone.

5.5.3 Utility of RS models and further study

While the combination of optical and radar imagery has been shown to be beneficial in calculating tropical forest biomass and distinguishing between temporal vegetation types (Fayad et al., 2016; Lopes et al., 2020), our findings indicated that models fitted with optical reflectance data underperformed. This is likely due to persistent cloud cover reducing image quality and thus the ability to capture fine scale changes in habitat (Nazarova et al., 2020). In the Northern Brazilian Amazon, for instance, Landsat has been shown to fail to deliver usable images for a minimum of 10 months per year (Quiñones et al., 2007). Additionally, forest loss has been shown to increase cloud cover in much of the central and northern Amazon (Xu et al., 2022). Despite the potential reduction in cloud-related limitations through the use of Planet satellites and their daily imagery, our results did not reflect this advantage, as the best performing RS model did not include information from an optical satellite. In contrast, the use of radar imagery, which is not affected by cloud cover, has demonstrated potential for improving the delineation between Boreal Forest types (Thiel et al., 2006). Radar imagery was the only satellite data included in the best-performing RS model. However, while radar is unaffected by cloud cover, variations in moisture content, both in and on leaves can influence backscatter, making comparisons between similar vegetation communities challenging across times or areas with differing water content (van Emmerik et al., 2017). The upcoming launch of a longer-wavelength radar satellite, NISAR (NASA-ISRO Synthetic Aperture Radar) may help overcome some of these challenges. Its longer wavelengths in comparison to Sentinel-1, should be less impacted by water content and better penetrate forest canopies to describe sub-canopy structures. This may prove particularly beneficial for studies conducted in high cloud areas, especially when considering phenology.

The limited range of predictive performance observed between models utilising different covariate types precludes the formulation of robust conclusions. However, this study presents an investigation into the potential causes and implications of these observed differences. The prevalence of cloud cover in tropical forests presents a significant challenge for the use of optical remote sensing in monitoring these ecosystems on a regular temporal basis. Meanwhile, models that performed the least well did not account for spatial autocorrelation. This reinforces the recommendation that, whatever other covariates are used, spatial factor models should be employed when prediction is a key goal (Doser et al., 2023).

While variations in bird communities could be effectively modelled at a categorical level, they struggled to capture finer-scale variations in bird community composition. This was likely due to the limited number of survey sites, which prevented effective training of models to detect gradual rather than broad habitat changes. This suggests that a substantially larger dataset, covering a broader range of habitat conditions would be needed for RS data to move beyond predicting gradient level variation to predicting fine-scale community shifts within those gradients (Wisz et al., 2008). The models indicate that there is a distinction in the diversity and likelihood of occurrence of bird species across forest gradients, and that there seems to be a correlation between forest gradients and RS data. In the context of our study, the mean detection level of numerous species was notably low, which has the potential to result in a considerable number of false absences, thereby exerting a detrimental impact on the subsequent data analysis. However, the high repeat, short sample acoustic survey technique employed, provides good levels of detection of dawn calling species. Improvements in detection and a more thorough community description may be attained by sampling acoustic recordings at different times of day to capture predominantly dusk and night calling species (Wimmer et al., 2013). While passive acoustic surveys provide an efficient non-invasive method of assessment, implementing complementary techniques such as mist netting and point surveys could improve the identification of low-vocalisation species (Martin et al., 2017). Furthermore, increasing the number of survey sites may facilitate analysis in two ways. Firstly, the provision of additional data will facilitate the confirmation of any observed variation in community richness and composition between forest gradients. Secondly, the incorporation of more RS variables in the analysis is possible with the inclusion of more data points, thus increasing the amount of variation in reflectance and structure that may be included in a model while avoiding the issue of overfitting. The incorporation of more RS variables allows for the inclusion of a greater range of variation in forest reflectance in analysis, potentially enhancing the capacity of models to discriminate and predict across gradients and landscapes. Should more training data improve the ability to model bird communities using RS data, it may be possible to accurately predict species richness and community similarities regionally and temporally, with minimal fieldwork. The findings from this study could help identify areas where shifts in bird community composition or species richness are occurring, aiding long-term biodiversity monitoring. Improving the predictive accuracy of RS models, particularly with larger datasets, could provide a scalable tool to track ecological changes and assess the effectiveness of management interventions over time.

Code and data used in the analysis of this chapter are available at
https://github.com/AndrewCSlater/Chapter_Brazil

6. General Discussion

6.1 Introduction and rationale

The loss and degradation of natural habitats is occurring at an alarming rate, with the result that species richness and diversity are declining across the world (IPBES, 2016; Newbold et al., 2016). This, in turn, impairs the capacity of ecosystems to sustain the services and functions they provide (Oliver et al., 2015). Furthermore, the effects of climate change, which are being exacerbated by deforestation, are intensifying the impacts of habitat degradation (IPCC, 2023; Malhi et al., 2009). It is of the utmost importance to halt and reverse these trends. The initial step in addressing this decline, as recognised in global biodiversity monitoring frameworks such as GEO BON (2023), is to describe the distribution and status of species, and to identify ecologically important environments and areas that are experiencing the most significant ecological decline. In order to have a meaningful impact on areas and species in decline, it is necessary to understand the ecological and anthropogenic drivers that influence change in them. These include processes such as habitat loss and fragmentation, climate change, overexploitation, invasive species, and pollution (IPBES, 2016; Maxwell et al., 2016; Newbold et al., 2016). Similarly, the effectiveness of conservation practices and the impact of landscape degradation must be quantified, not only currently but also by continued monitoring into the future. The data required to effectively describe a natural environment and changes within it often necessitate the use of spatial and temporal scales that exceed the practical capabilities of traditional *in situ* field surveys (Jetz et al., 2019; Schmeller et al., 2017). In response to these needs, there has been a growing trend in the use of remotely sensed satellite data to classify land use, to monitor its change, and to expand survey findings of a few limited species across landscapes. However, the classification process is inherently subjective and susceptible to error. The classes defined may not be pertinent to driving observed changes and may not reflect the true nature of the observed landscape (Foody, 2002; Olofsson et al., 2014). Similarly, obtaining information for individual species may aid in their protection, but even if they are keystone or flagship species, they are unlikely to fully describe the condition of an ecosystem. A more comprehensive picture of an environment and a more ecologically informative indicator of its condition or functioning, can be achieved by describing the composition and structure of species found within it. Community-level data reflect environmental gradients, species interactions, and disturbance regimes, and are widely used to assess ecosystem integrity and resilience (Andersen et al., 2002; Niemi and McDonald, 2004; Noss, 1990). This study advances the field by integrating multi-species modelling with remotely sensed data to assess whether improvements can be made to community predictions made by traditional environmental metrics. While Earth observation offers an invaluable tool for large-scale biodiversity assessments, its effectiveness is dependent on factors such as resolution, sensor type and environmental conditions (Pettorelli et al., 2014a; Turner et al., 2003; Wang and Gamon, 2019). These challenges highlight the need to determine where and when EO is most appropriate for predicting community composition.

6.2 Key findings across the studies

In Chapter Two, the hyper-diverse plant communities of the Greater Cape Floristic Region in South Africa were modelled. It was hypothesised that plant communities would directly influence the reflectance values measured by Earth observation satellites. However, it was found that the explanatory powers of distribution models fitted with geological data were only marginally enhanced by the incorporation of Earth observation data. Due to the greater influence that more abundant plants have on the reflectance values measured by satellites, the explanatory power of abundance models was greatly increased by Earth observation. The contribution of Earth observation to distribution models was minimal, and static topographic and climatic variables remained the strongest predictors. EO replaced a large proportion of the variance explained by latent spatial factors, and thereby accounted for a far greater proportion of the explained variance than it added. In contrast to latent variables, Earth observation is a directly measurable phenomenon, and Earth observation improved the ability to predict and map to areas outside of the training data. The resolution of satellite data permitted the identification of a greater diversity of community composition across the landscape than was apparent from traditional categorical vegetation maps. Furthermore, the use of historic satellite imagery permitted the prediction and mapping of areas where a directional shift in community composition may have occurred. This ability to detect temporal trends suggests that EO could be a valuable tool for long-term biodiversity monitoring, particularly in landscapes experiencing gradual shifts in species composition due to climate variability or land-use change.

In Chapter Three, insect communities surveyed in the community forest surrounding the Gola Rainforest National Park of Sierra Leone, and identified by DNA metabarcoding, were analysed. Models were able to accurately predict (AUC>0.7) the distribution of ~19% of the modelled insect OTUs, with Earth observation data accurately predicting slightly more OTUs (55) than habitat variables measured *in situ* could (49). The degree of overlap between the two groups of well-predicted insects was minimal, indicating that Earth observation data and habitat variables were measuring different elements of forest structure. Despite this, Earth observation was shown to explain a third of the variation in canopy structure as measured by GEDI, with many of the well-predicted OTUs exhibiting a negative correlation with canopy structure. Furthermore, GLCM-based EO variables outperformed raw reflectance values, reinforcing the importance of spatial texture in predicting biodiversity patterns and demonstrating how spatial heterogeneity in remotely sensed data better reflects forest condition. By measuring the richness of OTUs grouped by their positive or negative correlations with forest structure, forest condition could be inferred and mapped over the area. In addition, communities were mapped by their similarity of composition over the same area. These maps could be employed to identify areas where the intensification of farming would have the least detrimental impact on forest structure (Phalan et al., 2011). Furthermore, they could be employed as a means of measuring biodiversity, thus enabling the calculation of the success of REDD+ incentive payments, and the apportionment of such payments.

In Chapter Four, the bird communities in the Madre de Dios region of Peru were determined through multiple surveys and subsequently modelled. The results demonstrated that, once again, models incorporating data from Earth observation satellites performed better when predicting community composition than models with environmental variables, with 36% and 15% of modelled species respectively being accurately predicted (AUC>0.7). The overall species detectability was found to be very low, with variation in community composition across the landscape predicted to be far lower than observed, and despite the success in predicting individual species, communities were predicted poorly. To observe community composition more completely, the importance of greatly increasing the number of replicate surveys was emphasised, with recommendations that prioritising more replicate surveys at fewer sites could significantly improve detection and enhance model reliability.

In Chapter Five, the bird communities identified by passive acoustic monitoring in the northern Brazilian Amazon were modelled. EO models performed comparably to habitat variable models, predicting ~18% of species well (AUC>0.7), but their ability to capture fine-scale habitat differences was limited. Although the Earth observation data models performed adequately, the limited number of survey locations meant that fewer covariates could be included, requiring the data to be heavily condensed, thereby diluting the correlation between EO and survey data.

The four case studies highlight how the effectiveness of EO for biodiversity modelling is shaped by taxonomic group, survey design, and the spatial extent of sampling. In the GCFR, EO significantly improved plant abundance models, likely due to the direct relationship between vegetation structure and spectral reflectance. In contrast, EO struggled to predict bird community composition in Brazil and Peru, but for different reasons. In Peru, where survey coverage was extensive, the challenge arose because species detection was too low to adequately describe community variation across the area. In Brazil however, the primary limitation was the small number of sample sites, which restricted the models’ ability to differentiate habitat variation. The Gola study provided a more mixed result, where EO predictors explained variation for some taxa but not others, reinforcing that even within a single ecosystem, species respond differently to remotely sensed habitat features. These findings suggest that EO achieves the most accurate predictions of community composition when the environmental variables influencing species distributions, such as vegetation structure or land cover variability, can be reliably captured through satellite measurement. However, this is only part of the picture.

The design, scale, and ecological representativeness of the field data also had an influential role in shaping model outcomes. The GCFR was unique in spanning multiple biomes and transition zones, introducing the challenge that some vegetation communities may have shared similar spectral reflectance, but contained distinct plant assemblages, potentially limiting EO’s ability to distinguish them. In Brazil, the restricted number of sample sites constrained model performance, despite the relative completeness of the acoustic dataset. These findings emphasise that EO’s predictive power is strongest when applied at appropriate spatial scales and combined with well-structured field data that captures sufficient habitat variation.

6.3 Methodological challenges and reflections

A key methodological challenge addressed by this study was how to scale up species and community modelling using remotely sensed data across ecologically diverse systems. These challenges were particularly evident in the selection and reduction of EO predictors. While the method used to derive predictors through canonical or principal components may have incorporated more EO variables than necessary, this approach is not uncommon in other studies (Guo et al., 2023; Jarocińska et al., 2024; Tsai et al., 2007). However, a more refined strategy, selecting fewer but potentially more ecologically relevant indices for reduction, or using selected, unreduced spectral indices (Carlson et al., 2007; Rocchini et al., 2010; Xue and Su, 2017), might have improved model predictive performance. These methodological decisions were made in response to the need for continuous variables that could describe the landscape in a consistent way across diverse environments. The aim was to reduce subjectivity and the risk of incorrect classification, with the objective of enabling repeatable descriptions of community composition and species occurrence that could be mapped over areas of interest. While this approach allowed the study to explore the potential of EO at broad scales, it introduced trade-offs between ecological accuracy and model predictive strength.

It is important to consider the potential implications of spatial resolution when undertaking future studies. Firstly, it is important to consider what resolution is necessary and practically achievable. As new satellite platforms become available, image resolutions are becoming finer. Although not yet freely available, sub-1m^2 resolutions exist. It may be tempting to choose the highest resolution available, but this may not be necessary, as demonstrated in Chapter 3, where satellite platform, and thus resolution, had no significant influence when summarising reflectance over an area. Similarly, Sentinel-2 imagery with a 10m^2 resolution has nine pixels for every Landsat pixel, and a 1m^2 resolution image contains 100 pixels for every Sentinel pixel. While finer resolutions result in crisper images, the data storage required increases in proportion to the resolution. When analysing pixel-level data in complex statistical models, the additional data size necessitates increased requirements in processing power, time, memory and energy use. In the context of large-scale prediction, it is probable that the practical limitations of processing power will outweigh the potential benefits of smaller pixels. Secondly, the resolution of satellite pixels and field survey data should be as closely aligned as possible, particularly when modelling fine-scale ecological variation. However, the appropriate resolution may differ depending on the taxa being surveyed. For example, plant communities may respond to microhabitat variation at very fine scales, while birds often respond to broader habitat variations. Many field surveys in Chapter 2 were conducted on linear transects or in 10m^2 plots, yet the analysis used satellite data with a 30m^2 resolution. In these instances, the field data represented only a fraction of a single satellite pixel, diluting the relationship between the two. Moreover, the spatial resolution of EO data does not always correspond to the ecological scale at which species respond to environmental variation. Similarly, texture-based metrics such as GLCM, or Rao's Q (Rocchini et al., 2018), can provide additional information by describing heterogeneity within and between pixels. These approaches may capture relevant ecological variation at scales not directly resolved by raw EO data.

In predominantly cloud-covered areas, such as tropical regions, the use of optical EO data can be severely limited. Rather than concentrate on finer spatial resolution, which may not resolve cloud issues, studies in these areas may benefit from improved temporal resolution. More frequent image acquisition increases the likelihood of obtaining a cloud-free image, supporting the monitoring of short-term environmental variability. In addition, radar-based imagery is capable of penetrating cloud cover, and presents an increasingly viable alternative. New radar satellites, such as the upcoming BIOMASS mission, are anticipated to become operational in the near future. These platforms will offer the ability to penetrate both cloud cover and dense canopy cover and provide a more detailed description than is currently achievable (Quegan et al., 2019; Singh et al., 2019).

In addition to considerations of spatial and temporal resolution, survey design decisions can also influence model outcomes. For example, in passive acoustic monitoring, limiting analysis to a single 1-hour time window may introduce bias by over-representing species that vocalise frequently (Fairbrass et al., 2017), while under-sampling rarer or less vocal species. Future studies should evaluate whether extending the sampling duration improves species detectability and reduces biases (Bradfer-Lawrence et al., 2019), while also considering whether the potential improvements justify the additional resources required for analysing larger datasets (Williams et al., 2018).

6.4 Practical implications and conservation applications

Despite widespread international commitments to halt biodiversity loss, conservation outcomes remain well below targets, in part due to insufficient ecological data and poor alignment between monitoring systems and decision-making processes (CBD Secretariat, 2021; IPBES, 2016). The failure to meet the Aichi Biodiversity Targets has led to renewed calls for scalable, reliable biodiversity data to inform conservation planning (GEO BON, 2023; Mace et al., 2012). Monitoring frameworks now place growing emphasis on spatially explicit, cost-effective and repeatable tools for tracking biodiversity trends, particularly in regions with sparse field data (Jetz et al., 2012; Pettorelli et al., 2014a). The findings of this study contribute directly to these goals by demonstrating how EO data, in combination with ecological modelling, can expand the reach of field surveys and support biodiversity assessments over large spatial and temporal scales.

The models produced by this study provide a probability of occurrence for each species modelled at each point of prediction. In the search for rare or difficult-to-find species, conducting field surveys in areas with a higher chance of success can assist ecologists in increasing the efficiency of their surveys. The efficiency with which species of interest can be located is of particular benefit when identifying areas to focus conservation efforts. Greater efficiency should ideally enable more surveys to be conducted, which would in turn improve understanding of species distributions and provide better training data for new models.
Similarly, the results of a limited number of field surveys can be employed in the calculation of communities based on their compositional similarities and mapped across areas larger than

those surveyed. Such maps may be of benefit to policy makers in identifying areas of ecological significance, which can then be assigned a conservation status. Furthermore, these maps can be used to identify areas that are under threat, thereby advising where remedial action should be taken. However, the effectiveness of EO-based biodiversity models was highly dependent on ecosystem characteristics, data quality, and species detectability. While EO successfully captured broad-scale habitat gradients in some cases (Brazil, GRNP) and improved abundance models where species presence was confirmed (GCFR), it was less effective in predicting community composition in regions where detectability was low (Peru) (Kissling et al., 2015). Across all study sites, EO provided valuable insights for large-scale biodiversity monitoring, but its conservation applications depend on careful integration with ground-based surveys (Pettorelli et al., 2014a). The findings reinforce EO's role as a scalable, cost-effective tool for planning, habitat restoration, and tracking long-term ecological change (Turner et al., 2015). However, conservation practitioners should interpret model outputs with caution and supplement them with direct field validation where possible (Pettorelli et al., 2014b).

6.5 Limitations and future research

This study was based on secondary ecological data, gathered by various researchers for a range of independent purposes, and not originally designed with satellite remote sensing or this study in mind. As such, there was variability in the survey effort and methodologies employed, but we endeavoured to account for this within model structures. Although the primary objective of this project was to assess the efficacy of satellite data in describing field-based ecological surveys, the link between each chapter and ecological processes was contingent upon the environmental data gathered by the original survey and may not be the most appropriate or best fit. Despite these irregularities, the breadth of data made available was far greater than could be gathered by an individual PhD study, and as such were as much a benefit as limitation. Finally, they can be employed to promote sustainable development of local communities by identifying areas best suited for development with minimal impact on the ecology of the local environment.

6.6 Conclusion

In all cases, models fitted with remotely sensed satellite data were found to be as effective, or even more so, at predicting species distributions than models fitted with environmental data gathered *in situ*. However, the effectiveness of EO-based models varied across ecosystems and taxa, with greater accuracy in predicting species abundance in well surveyed environments and broad habitat gradients, but lower predictive success was achieved in landscapes with high species turnover and limitations in detectability. This evidence suggests that the use of satellite data can be effectively advanced from the prediction of land use or large, prevalent, plant species, to the collective description of bird, insect, or small plant communities (Pettorelli et al., 2014a).

Our study also enables the inference of community composition in the past using historical satellite data, and the continual calculation of this into the future with each new satellite image produced. As a result, land managers are able to ascertain the consequences of historical practices and observe the impact of degradation and the efficacy of ongoing mitigation strategies on the land under their stewardship.

It is necessary to be able to describe both the likely distribution of individual species and the composition of ecological communities over space and time. The findings from across all four study systems indicate that EO can reliably capture broad-scale biodiversity patterns, but its ability to predict finer-scale ecological interactions remains dependant on species detectability, survey effort, and ecosystem structure (Turner et al., 2015). The utilisation of satellite data to forecast community composition over extensive geographical areas is a viable approach in a variety of environments and for a diverse range of taxa. Furthermore, joint species distribution and occupancy models, which use satellite-derived variables, are generally able to predict ecological communities with greater accuracy than those fitted with environmental variables described *in situ* (Pollock et al., 2014). Future research should focus on refining model calibration methods, improving integration with *in situ* data, and assessing the long-term stability of EO based biodiversity predictions (Pettorelli et al., 2014a).

The output of models enables the mapping of species or communities, thereby facilitating the inference of ecological condition over extended areas of interest. This should benefit ecologists, land managers, policy makers, and human development agencies in achieving their goals. By synthesising results across multiple ecosystems , this study demonstrates that EO has the potential to complement traditional biodiversity surveys and expand ecological monitoring to previously inaccessible areas, reinforcing its role in large-scale conservation planning. This provides further evidence that biodiversity surveys can be successfully linked with and expanded by Earth observation data, and should facilitate future studies to expand into areas, taxa, and ecological variables where ecological surveys are currently limited (Kissling et al., 2015).

References

Alvarez-Berríos, N., Campos-Cerqueira, M., Hernández-Serna, A., Amanda Delgado, C.J., Román-Dañobeytia, F., Aide, T.M., 2016. Impacts of Small-Scale Gold Mining on Birds and Anurans Near the Tambopata Natural Reserve, Peru, Assessed Using Passive Acoustic Monitoring. Trop Conserv Sci 9, 832–851. https://doi.org/10.1177/194008291600900216

Andersen, A.N., Hoffmann, B.D., Müller, W.J., Griffiths, A.D., 2002. Using ants as bioindicators in land management: simplifying assessment of ant community responses. Journal of Applied Ecology 39, 8–17. https://doi.org/10.1046/j.1365-2664.2002.00704.x

Anderson, D.L., 2009. Ground Versus Canopy Methods for the Study of Birds in Tropical Forest Canopies: Implications for Ecology and Conservation. Condor 111, 226–237. https://doi.org/10.1525/cond.2009.090032

Anderson, R.P., Raza, A., 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus Nephelomys) in Venezuela. J Biogeogr 37, 1378–1393. https://doi.org/10.1111/j.1365-2699.2010.02290.x

Asner, G.P., 2007. Carnegie Airborne Observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. J Appl Remote Sens 1, 013536. https://doi.org/10.1117/1.2794018

Asner, G.P., Kellner, J.R., Kennedy-Bowdoin, T., Knapp, D.E., Anderson, C., Martin, R.E., 2013. Forest Canopy Gap Distributions in the Southern Peruvian Amazon. PLoS One 8, e60875. https://doi.org/10.1371/journal.pone.0060875

Asner, G.P., Powell, G.V.N., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M., Hughes, R.F., 2010. High-resolution forest carbon stocks and emissions in the Amazon. Proceedings of the National Academy of Sciences 107, 16738–16742. https://doi.org/10.1073/pnas.1004875107

Aubry, K.B., Raley, C.M., McKelvey, K.S., 2017. The importance of data quality for generating reliable distribution models for rare, elusive, and cryptic species. PLoS One 12, e0179152. https://doi.org/10.1371/journal.pone.0179152

Azzari, G., Lobell, D.B., 2017. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. Remote Sens Environ 202, 64–74. https://doi.org/10.1016/j.rse.2017.05.025

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R. Mac, Thomson, J.R., Ferraz, S.F. de B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro de Castro Solar, R., Vieira, I.C.G., Aragão, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., de Oliveira, R.C., Souza Jr, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-de-Mello, F.Z., Veiga, R.C.S., Venturieri, A., Gardner, T.A., 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature 535, 144–147. https://doi.org/10.1038/nature18326

Barlow, J., Mestre, L.A.M., Gardner, T.A., Peres, C.A., 2007. The value of primary, secondary and plantation forests for Amazonian birds. Biol Conserv 136, 212–231. https://doi.org/10.1016/j.biocon.2006.11.021

Barlow, J., Peres, C.A., Henriques, L.M.P., Stouffer, P.C., Wunderle, J.M., 2006. The responses of understorey birds to forest fragmentation, logging and wildfires: An Amazonian synthesis. Biol Conserv 128, 182–192. https://doi.org/10.1016/j.biocon.2005.09.028

Barros, F. de C., Almeida, S.M., Godoy, B.S., Silva, R.R. da, Silva, L.C., de Moraes, K.F., Santos, M.P.D., 2022. Taxonomic and functional diversity of bird communities in mining areas undergoing passive and active restoration in eastern Amazon. Ecol Eng 182, 106721. https://doi.org/10.1016/j.ecoleng.2022.106721

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Basile, M., Storch, I., Mikusiński, G., 2021. Abundance, species richness and diversity of forest bird assemblages – The relative importance of habitat structures and landscape context. Ecol Indic 133, 108402. https://doi.org/10.1016/j.ecolind.2021.108402

Basset, Y., Cizek, L., Cuénoud, P., Didham, R.K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A.K., Winchester, N.N., Roubik, D.W., Aberlenc, H.-P., Bail, J., Barrios, H., Bridle, J.R., Castaño-Meneses, G., Corbara, B., Curletti, G., Duarte da Rocha, W., De Bakker, D., Delabie, J.H.C., Dejean, A., Fagan, L.L., Floren, A., Kitching, R.L., Medianero, E., Miller, S.E., Gama de Oliveira, E., Orivel, J., Pollet, M., Rapp, M., Ribeiro, S.P., Roisin, Y., Schmidt, J.B., Sørensen, L., Leponce, M., 2012. Arthropod Diversity in a Tropical Forest. Science (1979) 338, 1481–1484. https://doi.org/10.1126/science.1226727

Bell, W., Hoffman, M.T., Visser, V., Kirsten, T., 2023. Modelling land condition to augment Land Degradation Neutrality assessments - The succulent Karoo biome of South Africa as a case study. J Arid Environ 219, 105086. https://doi.org/10.1016/j.jaridenv.2023.105086

Betts, M.G., Yang, Z., Hadley, A.S., Hightower, J., Hua, F., Lindenmayer, D., Seo, E., Healey, S.P., 2024. Quantifying forest degradation requires a long-term, landscape-scale approach. Nat Ecol Evol. https://doi.org/10.1038/s41559-024-02409-5

Blackburn, G.A., Milton, E.J., 1995. Seasonal variations in the spectral reflectance of deciduous tree canopies. Int J Remote Sens 16, 709–720. https://doi.org/10.1080/01431169508954435

Blake, J.G., Loiselle, B.A., 2015. Enigmatic declines in bird numbers in lowland forest of eastern Ecuador may be a consequence of climate change. PeerJ 3, e1177. https://doi.org/10.7717/peerj.1177

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W., de Bruyn, M., 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol Evol 29, 358–367. https://doi.org/10.1016/j.tree.2014.04.003

Bolam, F.C., Mair, L., Angelico, M., Brooks, T.M., Burgman, M., Hermes, C., Hoffmann, M., Martin, R.W., McGowan, P.J.K., Rodrigues, A.S.L., Rondinini, C., Westrip, J.R.S., Wheatley, H., Bedolla-Guzmán, Y., Calzada, J., Child, M.F., Cranswick, P.A., Dickman, C.R., Fessl, B., Fisher, D.O., Garnett, S.T., Groombridge, J.J., Johnson, C.N., Kennerley, R.J., King, S.R.B., Lamoreux, J.F., Lees, A.C., Lens, L., Mahood, S.P., Mallon, D.P., Meijaard, E., Méndez-Sánchez, F., Percequillo, A.R., Regan, T.J., Renjifo, L.M., Rivers, M.C., Roach, N.S., Roxburgh, L., Safford, R.J., Salaman, P., Squires, T., Vázquez-Domínguez, E., Visconti, P., Woinarski, J.C.Z., Young, R.P., Butchart, S.H.M., 2021. How many bird and mammal extinctions has recent conservation action prevented? Conserv Lett 14, e12762. https://doi.org/10.1111/conl.12762

Born, J., Linder, H.P., Desmet, P., 2007. ORIGINAL ARTICLE: The Greater Cape Floristic Region. J Biogeogr 34, 147–162. https://doi.org/10.1111/j.1365-2699.2006.01595.x

Boucher, F.C., Verboom, G.A., Musker, S., Ellis, A.G., 2017. Plant size: a key determinant of diversification? New Phytologist 216, 24–31. https://doi.org/10.1111/nph.14697

Bourgoin, C., Blanc, L., Bailly, J.-S., Cornu, G., Berenguer, E., Oszwald, J., Tritsch, I., Laurent, F., Hasan, A.F., Sist, P., Gond, V., 2018. The Potential of Multisource Remote Sensing for Mapping the Biomass of a Degraded Amazonian Forest. Forests 9, 303. https://doi.org/10.3390/f9060303

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S.G., Dent, D.H., 2019. Guidelines for the use of acoustic indices in environmental research. Methods Ecol Evol 10, 1796–1807. https://doi.org/10.1111/2041-210X.13254

Bradshaw, C.J., Sodhi, N.S., Brook, B.W., 2009. Tropical turmoil: a biodiversity tragedy in progress. Front Ecol Environ 7, 79–87. https://doi.org/10.1890/070193

Bregman, T.P., Lees, A.C., Seddon, N., MacGregor, H.E.A., Darski, B., Aleixo, A., Bonsall, M.B., Tobias, J.A., 2015. Species interactions regulate the collapse of biodiversity and ecosystem function in tropical forest fragments. Ecology 96, 2692–2704. https://doi.org/10.1890/14-1731.1

Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Czerwinski, W., Pasquarella, V.J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., Tait, A.M., 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. Sci Data 9, 251. https://doi.org/10.1038/s41597-022-01307-4

Buchner, D., Beermann, A., Hörren, T., Enss, J., Frenzel, M., Li, Y., Müller, J., Pauls, S.U., Sorg, M., consortium, L.-D., Haase, P., Leese, F., 2023. German-wide Malaise trap metabarcoding estimates over 33,000 insect species. bioRxiv 2023.05.04.539402. https://doi.org/10.1101/2023.05.04.539402

Burivalova, Z., Lee, T.M., Giam, X., Şekercioğlu, Ç.H., Wilcove, D.S., Koh, L.P., 2015. Avian responses to selective logging shaped by species traits and logging practices. Proceedings of the Royal Society B: Biological Sciences 282, 20150164. https://doi.org/10.1098/rspb.2015.0164

Burton, C., Betts, R.A., Jones, C.D., Feldpausch, T.R., Cardoso, M., Anderson, L.O., 2020. El Niño Driven Changes in Global Fire 2015/16. Front Earth Sci (Lausanne) 8. https://doi.org/10.3389/feart.2020.00199

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., Martius, C., Zlinszky, A., Calvignac-Spencer, S., Cobbold, C.A., Dawson, T.P., Emerson, B.C., Ferrier, S., Gilbert, M.T.P., Herold, M., Jones, L., Leendertz, F.H., Matthews, L., Millington, J.D.A., Olson, J.R., Ovaskainen, O., Raffaelli, D., Reeve, R., Rödel, M.O., Rodgers, T.W., Snape, S., Visseren-Hamakers, I., Vogler, A.P., White, P.C.L., Wooster, M.J., Yu, D.W., 2017. Connecting Earth observation to high-throughput biodiversity data. Nat Ecol Evol 1, 1–9. https://doi.org/10.1038/s41559-017-0176

Caddy-Retalic, S., Wardle, G.M., Leitch, E.J., Mcinerney, F.A., Lowe, A.J., 2020. Vegetation change along a Mediterranean to arid zone bioclimatic gradient reveals scale-dependent ecotone patterning. Aust J Bot 68, 574–586. https://doi.org/10.1071/BT20036

Cagnina, C., Marino, A., Silva-Perez, C., Ruiz-Ramos, J., Suarez, J., 2023. Assessment of the Impact of Surface Water Content for Temperate Forests in SAR Data at C-Band. Remote Sens (Basel) 15, 5723. https://doi.org/10.3390/rs15245723

CapeNature, 2025. Corridors - Capenature [WWW Document]. URL https://www.capenature.co.za/corridors (accessed 1.20.25).

Carlson, K.M., Asner, G.P., Hughes, R.F., Ostertag, R., Martin, R.E., 2007. Hyperspectral Remote Sensing of Canopy Biodiversity in Hawaiian Lowland Rainforests. Ecosystems 10, 536–549. https://doi.org/10.1007/s10021-007-9041-z

Carvalho, W.D., Fluck, I.E., de Castro, I.J., Hilário, R.R., Martins, A.C.M., de Toledo, J.J., da Silva Xavier, B., Dambros, C., Bobrowiec, P.E.D., 2023. Elevation drives taxonomic, functional and phylogenetic β-diversity of phyllostomid bats in the Amazon biome. J Biogeogr 50, 70–85. https://doi.org/10.1111/jbi.14533

CBD, 2010. Aichi Biodiversity Targets. Aichi Bidiversity Targets 9–10.

CBD Secretariat, 2021. First Draft of the Post-2020 Global Biodiversity Framework, Cbd/Wg2020/3/3.

Ceballos, G., Ehrlich, P.R., 2002. Mammal Population Losses and the Extinction Crisis. Science (1979) 296, 904–907. https://doi.org/10.1126/science.1069349

CEPF, 2009. Succulent Karoo | CEPF [WWW Document]. URL https://www.cepf.net/our-work/biodiversity-hotspots/succulent-karoo (accessed 1.20.25).

Chalmandrier, L., Stouffer, D.B., Purcell, A.S.T., Lee, W.G., Tanentzap, A.J., Laughlin, D.C., 2022. Predictions of biodiversity are improved by integrating trait-based competition with abiotic filtering. Ecol Lett 25, 1277–1289. https://doi.org/10.1111/ele.13980

Chauvier, Y., Thuiller, W., Brun, P., Lavergne, S., Descombes, P., Karger, D.N., Renaud, J., Zimmermann, N.E., 2021. Influence of climate, soil, and land cover on plant species distribution in the European Alps. Ecol Monogr 91. https://doi.org/10.1002/ecm.1433

Chazdon, R.L., Brancalion, P.H.S., Laestadius, L., Bennett-Curry, A., Buckingham, K., Kumar, C., Moll-Rocek, J., Vieira, I.C.G., Wilson, S.J., 2016. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. Ambio 45, 538–550. https://doi.org/10.1007/s13280-016-0772-y

Chuvieco, E., Riaño, D., Aguado, I., Cocero, D., 2002. Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: Applications in fire danger assessment. Int J Remote Sens 23, 2145–2162. https://doi.org/10.1080/01431160110069818

Clark, D.B., Oberbauer, S.F., Clark, D.A., Ryan, M.G., Dubayah, R.O., 2021. Physical structure and biological composition of canopies in tropical secondary and old-growth forests. PLoS One 16, e0256571. https://doi.org/10.1371/journal.pone.0256571

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S. V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens Environ 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002

Conservation International, 2023. Biodiversity Hotspots [WWW Document]. URL https://www.conservation.org/priorities/biodiversity-hotspots (accessed 6.14.23).

Cruz, M., Pradel, W., Juarez, H., Hualla, V., Suarez, V., 2023. Deforestation Dynamics in Peru. A Comprehensive Review of Land Use, Food Systems, and Socio-Economic Drivers. https://doi.org/10.4160/CIP.2023.12.007

Csillik, O., Kumar, P., Mascaro, J., O’Shea, T., Asner, G.P., 2019. Monitoring tropical forest carbon stocks and emissions using Planet satellite data. Sci Rep 9, 17831. https://doi.org/10.1038/s41598-019-54386-6

Dapporto, L., Ramazzotti, M., Fattorini, S., Vila, R., Talavera, G., Dennis, R.H.L., 2020. recluster: Ordination Methods for the Analysis of Beta-Diversity Indices.

Dayaram, A., Harris, L.R., Grobler, B.A., Van der Merwe, S., Rebelo, A.G., Ward Powrie, L., Vlok, J.H.J., Desmet, P.G., Qabaqaba, M., Hlahane, K.M., Skowno, A.L., 2019. Vegetation

Map of South Africa, Lesotho and Swaziland 2018: A description of changes since 2006. Bothalia 49. https://doi.org/10.4102/abc.v49i1.2452

de la Fuente, A., Hirsch, B.T., Cernusak, L.A., Williams, S.E., 2021. Predicting species abundance by implementing the ecological niche theory. Ecography 44, 1723–1730. https://doi.org/10.1111/ecog.05776

Dias, A., Van Houdt, S., Meschin, K., Von Stackelberg, K., Bago, M.-L., Baldarelli, L., Gonzalez Downs, K., Luuk, M., Delubac, T., Bottagisio, E., Kasak, K., Kebabci, A., Levers, O., Miilvee, I., Paju-Hamburg, J., Poncet, R., Sanfilippo, M., Sildam, J., Stepanov, D., Karnauskaite, D., 2023. Using essential biodiversity variables to assess forest ecosystem integrity. Frontiers in Forests and Global Change 6, 1098901. https://doi.org/10.3389/ffgc.2023.1098901

Dirzo, R., Young, H.S., Galetti, M., Ceballos, G., Isaac, N.J.B., Collen, B., 2014. Defaunation in the Anthropocene. Science (1979) 345, 401–406. https://doi.org/10.1126/science.1251817

Dorigo, W., Lucieer, A., Podobnikar, T., Čarni, A., 2012. Mapping invasive Fallopia japonica by combined spectral, spatial, and temporal analysis of digital orthophotos. International Journal of Applied Earth Observation and Geoinformation 19, 185–195. https://doi.org/10.1016/j.jag.2012.05.004

Dormann, C.F., 2007. Promising the future? Global change projections of species distributions. Basic Appl Ecol 8, 387–397. https://doi.org/10.1016/j.baae.2006.11.001

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Doser, J., 2023. Convergence diagnostics and other considerations when fitting spatial occupancy models [WWW Document]. https://www.jeffdoser.com/files/spoccupancy-web/articles/modelconsiderations.

Doser, J.W., Finley, A.O., Banerjee, S., 2023. Joint species distribution models with imperfect detection for high-dimensional spatial data. Ecology 104. https://doi.org/10.1002/ecy.4137

Doser, J.W., Finley, A.O., Kéry, M., Zipkin, E.F., 2022. spOccupancy: An R package for single-species, multi-species, and integrated spatial occupancy models. Methods Ecol Evol 13, 1670–1678. https://doi.org/10.1111/2041-210X.13897

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sens Environ 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Dubayah, R., Armston, J., Healey, S.P., Bruening, J.M., Patterson, P.L., Kellner, J.R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J.B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Hurtt, G., Luthcke, S., 2022. GEDI launches a new era of biomass inference from space. Environmental Research Letters 17, 095001. https://doi.org/10.1088/1748-9326/AC8694

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing 1, 100002. https://doi.org/10.1016/j.srs.2020.100002

Dunn, R.R., 2005. Modern Insect Extinctions, the Neglected Majority. Conservation Biology 19, 1030–1036. https://doi.org/10.1111/j.1523-1739.2005.00078.x

Dvořáková, D., Šipoš, J., Suchomel, J., 2023. Impact of agricultural landscape structure on the patterns of bird species diversity at a regional scale. Avian Res 14, 100147. https://doi.org/10.1016/j.avrs.2023.100147

Elias, F., Ferreira, J., Lennox, G.D., Berenguer, E., Ferreira, S., Schwartz, G., Melo, L. de O., Reis Júnior, D.N., Nascimento, R.O., Ferreira, F.N., Espirito-Santo, F., Smith, C.C., Barlow, J., 2020. Assessing the growth and climate sensitivity of secondary forests in highly deforested Amazonian landscapes. Ecology 101. https://doi.org/10.1002/ecy.2954

Ellis, A.G., Weis, A.E., Gaut, B.S., 2006. Evolutionary radiation of "Stone Plants" in the genus Argyroderma (Aizoaceae): Unravelling the effects of landscape, habitat, and flowering time. Evolution (N Y) 60, 39–55. https://doi.org/10.1111/j.0014-3820.2006.tb01080.x

Erwin, T.L., 1982. Tropical Forests: Their Richness in Coleoptera and Other Arthropod Species. Coleopt Bull 36, 74–75. https://doi.org/https://www.jstor.org/stable/4007977

Esquivel-Muelbert, A., Baker, T.R., Dexter, K.G., Lewis, S.L., Brienen, R.J.W., Feldpausch, T.R., Lloyd, J., Monteagudo-Mendoza, A., Arroyo, L., Álvarez-Dávila, E., Higuchi, N., Marimon, B.S., Marimon-Junior, B.H., Silveira, M., Vilanova, E., Gloor, E., Malhi, Y., Chave, J., Barlow, J., Bonal, D., Davila Cardozo, N., Erwin, T., Fauset, S., Hérault, B., Laurance, S., Poorter, L., Qie, L., Stahl, C., Sullivan, M.J.P., ter Steege, H., Vos, V.A., Zuidema, P.A., Almeida, E., Almeida de Oliveira, E., Andrade, A., Vieira, S.A., Aragão, L., Araujo-Murakami, A., Arets, E., Aymard C, G.A., Baraloto, C., Camargo, P.B., Barroso, J.G., Bongers, F., Boot, R., Camargo, J.L., Castro, W., Chama Moscoso, V., Comiskey, J., Cornejo Valverde, F., Lola da Costa, A.C., del Aguila Pasquel, J., Di Fiore, A., Fernanda Duque, L., Elias, F., Engel, J., Flores Llampazo, G., Galbraith, D., Herrera Fernández, R., Honorio Coronado, E., Hubau, W., Jimenez-Rojas, E., Lima, A.J.N., Umetsu, R.K., Laurance, W., Lopez-Gonzalez, G., Lovejoy, T., Aurelio Melo Cruz, O., Morandi, P.S., Neill, D., Núñez Vargas, P., Pallqui Camacho, N.C., Parada Gutierrez, A., Pardo, G., Peacock, J., Peña-Claros, M., Peñuela-Mora, M.C., Petronelli, P., Pickavance, G.C., Pitman, N., Prieto, A., Quesada, C., Ramírez-Angulo, H., Réjou-Méchain, M., Restrepo Correa, Z., Roopsind, A., Rudas, A., Salomão, R., Silva, N., Silva Espejo, J., Singh, J., Stropp, J., Terborgh, J., Thomas, R., Toledo, M., Torres-Lezama, A., Valenzuela Gamarra, L., van de Meer, P.J., van der Heijden, G., van der Hout, P., Vasquez Martinez, R., Vela, C., Vieira, I.C.G., Phillips, O.L., 2019. Compositional response of Amazon forests to climate change. Glob Chang Biol 25, 39–56. https://doi.org/10.1111/gcb.14413

European Commission, 2022. Corporate Sustainability Reporting Directive 2022/2464 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Official Journal of the European Union.

Fairbrass, A.J., Rennert, P., Williams, C., Titheridge, H., Jones, K.E., 2017. Biases of acoustic indices measuring biodiversity in urban areas. Ecol Indic 83, 169–177. https://doi.org/10.1016/j.ecolind.2017.07.064

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D.E., 2007. The shuttle radar topography mission. Reviews of Geophysics 45, RG2004. https://doi.org/10.1029/2005RG000183

Fayad, I., Baghdadi, N., Guitet, S., Bailly, J.-S., Hérault, B., Gond, V., El Hajj, M., Tong Minh, D.H., 2016. Aboveground biomass mapping in French Guiana by combining remote sensing, forest inventories and environmental data. International Journal of Applied Earth Observation and Geoinformation 52, 502–514. https://doi.org/10.1016/j.jag.2016.07.015

Feeley, K.J., Silman, M.R., 2011. Keep collecting: accurate species distribution modelling requires more collections than previously thought. Divers Distrib 17, 1132–1140. https://doi.org/10.1111/J.1472-4642.2011.00813.X

Ferrarini, A., Bai, Y., Dai, J., Alatalo, J.M., 2021. A new method for broad-scale modeling and projection of plant assemblages under climatic, biotic, and environmental cofiltering. Conservation Biology cobi.13797. https://doi.org/10.1111/cobi.13797

Ferraz, G., Nichols, J.D., Hines, J.E., Stouffer, P.C., Bierregaard, R.O., Lovejoy, T.E., 2007. A Large-Scale Deforestation Experiment: Effects of Patch Area and Isolation on Amazon Birds. Science (1979) 315, 238–241. https://doi.org/10.1126/science.1133097

Fiandino, S., Plevich, J., Tarico, J., Utello, M., Demaestri, M., Gyenge, J., 2020. Modeling forest site productivity using climate data and topographic imagery in Pinus elliottii plantations of central Argentina. Ann For Sci 77, 1–9. https://doi.org/10.1007/S13595-020-01006-3/FIGURES/3

Flores, B.M., Montoya, E., Sakschewski, B., Nascimento, N., Staal, A., Betts, R.A., Levis, C., Lapola, D.M., Esquível-Muelbert, A., Jakovac, C., Nobre, C.A., Oliveira, R.S., Borma, L.S., Nian, D., Boers, N., Hecht, S.B., ter Steege, H., Arieira, J., Lucas, I.L., Berenguer, E., Marengo, J.A., Gatti, L. V., Mattos, C.R.C., Hirota, M., 2024. Critical transitions in the Amazon forest system. Nature 626, 555–564. https://doi.org/10.1038/s41586-023-06970-0

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens Environ 80, 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

França, F.M., Frazão, F.S., Korasaki, V., Louzada, J., Barlow, J., 2017. Identifying thresholds of logging intensity on dung beetle communities to improve the sustainable management of Amazonian tropical forests. Biol Conserv 216, 115–122. https://doi.org/10.1016/j.biocon.2017.10.014

Frye, H.A., Aiello-Lammens, M.E., Euston-Brown, D., Jones, C.S., Kilroy Mollmann, H., Merow, C., Slingsby, J.A., van der Merwe, H., Wilson, A.M., Silander, J.A., 2021. Plant spectral diversity as a surrogate for species, functional and phylogenetic diversity across a hyper-diverse biogeographic region. Global Ecology and Biogeography 30, 1403–1417. https://doi.org/10.1111/geb.13306

Gallardo-Cruz, J.A., Solórzano, J. V., González, E.J., Meave, J.A., 2024. The Effect of Spatial Scale on the Prediction of Tropical Forest Attributes from Image Texture. International Journal of Forestry Research 2024, 1–10. https://doi.org/10.1155/2024/7178211

Gardner, T.A., Ferreira, J., Barlow, J., Lees, A.C., Parry, L., Vieira, I.C.G., Berenguer, E., Abramovay, R., Aleixo, A., Andretti, C., Aragão, L.E.O.C., Araújo, I., de Ávila, W.S., Bardgett, R.D., Batistella, M., Begotti, R.A., Beldini, T., de Blas, D.E., Braga, R.F., Braga, D. de L., de Brito, J.G., de Camargo, P.B., Campos dos Santos, F., de Oliveira, V.C., Cordeiro, A.C.N., Cardoso, T.M., de Carvalho, D.R., Castelani, S.A., Chaul, J.C.M., Cerri, C.E., Costa, F. de A., da Costa, C.D.F., Coudel, E., Coutinho, A.C., Cunha, D., D'Antona, Á., Dezincourt, J., Dias-Silva, K., Durigan, M., Esquerdo, J.C.D.M., Feres, J., Ferraz, S.F. de B., Ferreira, A.E. de M., Fiorini, A.C., da Silva, L.V.F., Frazão, F.S., Garrett, R., Gomes, A. dos S., Gonçalves, K. da S., Guerrero, J.B., Hamada, N., Hughes, R.M., Igliori, D.C., Jesus, E. da C., Juen, L., Junior, M., Junior, J.M.B. de O., Junior, R.C. de O., Junior, C.S., Kaufmann, P., Korasaki, V., Leal, C.G., Leitão, R., Lima, N., Almeida, M. de F.L., Lourival, R., Louzada, J., Nally, R. Mac, Marchand, S., Maués, M.M., Moreira, F.M.S., Morsello, C., Moura, N., Nessimian, J., Nunes, S., Oliveira, V.H.F., Pardini, R., Pereira, H.C., Pompeu, P.S., Ribas, C.R., Rossetti, F., Schmidt, F.A., da Silva, R., da Silva, R.C.V.M., da Silva, T.F.M.R., Silveira, J., Siqueira, J.V., de Carvalho, T.S., Solar, R.R.C., Tancredi, N.S.H., Thomson, J.R., Torres, P.C., Vaz-de-Mello, F.Z., Veiga, R.C.S., Venturieri, A., Viana, C., Weinhold, D., Zanetti, R., Zuanon, J., 2013. A social and ecological assessment of tropical land uses at multiple scales: the Sustainable Amazon Network. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20120166. https://doi.org/10.1098/rstb.2012.0166

Gelfand, A.E., 2020. Statistical challenges in spatial analysis of plant ecology data. Spat Stat 37, 100418. https://doi.org/10.1016/j.spasta.2020.100418

GEO BON, 2023. Strategic Plan 2023-2026.

GEO BON, 2018. What are EBVs? – GEO BON [WWW Document]. What are EBVs. URL https://geobon.org/ebvs/what-are-ebvs/ (accessed 7.22.22).

Gerlach, J., Samways, M., Pryke, J., 2013. Terrestrial invertebrates as bioindicators: an overview of available taxonomic groups. J Insect Conserv 17, 831–850. https://doi.org/10.1007/s10841-013-9565-9

Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A., Barlow, J., Peres, C.A., Bradshaw, C.J.A., Laurance, W.F., Lovejoy, T.E., Sodhi, N.S., 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. Nature 478, 378–381. https://doi.org/10.1038/nature10425

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gould, M.J., Gould, W.R., Cain, J.W., Roemer, G.W., 2019. Validating the performance of occupancy models for estimating habitat use and predicting the distribution of highly-mobile species: A case study using the American black bear. Biol Conserv 234, 28–36. https://doi.org/10.1016/j.biocon.2019.03.010

Guerin, G.R., Andersen, A.N., Rossetto, M., van Leeuwen, S., Byrne, M., Sparrow, B., Rodrigo, M., Lowe, A.J., 2019. Consistent sorting but contrasting transition zones in plant communities along bioclimatic gradients. Acta Oecologica 95, 74–85. https://doi.org/10.1016/j.actao.2019.01.006

Guillera-Arroita, G., Lahoz-Monfort, J.J., 2012. Designing studies to detect differences in species occupancy: power analysis under imperfect detection. Methods Ecol Evol 3, 860–869. https://doi.org/10.1111/j.2041-210X.2012.00225.x

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol Modell 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Guizar-Coutiño, A., Jones, J.P.G., Balmford, A., Carmenta, R., Coomes, D.A., 2022. A global evaluation of the effectiveness of voluntary REDD+ projects at reducing deforestation and degradation in the moist tropics. Conservation Biology 36, e13970. https://doi.org/10.1111/cobi.13970

Guo, Y., Mokany, K., Ong, C., Moghadam, P., Ferrier, S., Levick, S.R., 2023. Plant species richness prediction from DESIS hyperspectral data: A comparison study on feature extraction procedures and regression models. ISPRS Journal of Photogrammetry and Remote Sensing 196, 120–133. https://doi.org/10.1016/j.isprsjprs.2022.12.028

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens Environ 90, 337–352. https://doi.org/10.1016/j.rse.2003.12.013

Haddad, N.M., Brudvig, L.A., Clobert, J., Davies, K.F., Gonzalez, A., Holt, R.D., Lovejoy, T.E., Sexton, J.O., Austin, M.P., Collins, C.D., Cook, W.M., Damschen, E.I., Ewers, R.M., Foster, B.L., Jenkins, C.N., King, A.J., Laurance, W.F., Levey, D.J., Margules, C.R., Melbourne, B.A., Nicholls, A.O., Orrock, J.L., Song, D.-X., Townshend, J.R., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. Sci Adv 1. https://doi.org/10.1126/sciadv.1500052

Hall-Beyer, M., 2017. Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. Int J Remote Sens 38, 1312–1338. https://doi.org/10.1080/01431161.2016.1278314

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 (1979) 342, 850–853. https://doi.org/10.1126/science.1244693

Haralick, R.M., Dinstein, I., Shanmugam, K., 1973. Textural Features for Image Classification. IEEE Trans Syst Man Cybern SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Hardisky, M.A., Klemas, V., Smart, R.M., 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogramm Eng Remote Sensing 49, 77–83.

Harrison, M.E., Boonman, A., Cheyne, S.M., Husson, S.J., Marchant, N.C., Struebig, M.J., 2012. Biodiversity Monitoring Protocols for REDD+: Can a One-Size-Fits-All Approach Really Work? Trop Conserv Sci 5, 1–11. https://doi.org/10.1177/194008291200500102

He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.N., Schmidtlein, S., Turner, W., Wegmann, M., Pettorelli, N., 2015. Will remote sensing shape the next generation of species distribution models? Remote Sens Ecol Conserv 1, 4–18. https://doi.org/10.1002/rse2.7

Helfenstein, I.S., Schneider, F.D., Schaepman, M.E., Morsdorf, F., 2022. Assessing biodiversity from space: Impact of spatial and spectral resolution on trait-based functional diversity. Remote Sens Environ 275, 113024. https://doi.org/10.1016/j.rse.2022.113024

Hess, L.L., Melack, J.M., Affonso, A.G., Barbosa, C., Gastil-Buhl, M., Novo, E.M.L.M., 2015. Wetlands of the Lowland Amazon Basin: Extent, Vegetative Cover, and Dual-season Inundated Area as Mapped with JERS-1 Synthetic Aperture Radar. Wetlands 35, 745–756. https://doi.org/10.1007/s13157-015-0666-y

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965–1978. https://doi.org/10.1002/joc.1276

Hilker, T., Lyapustin, A.I., Tucker, C.J., Hall, F.G., Myneni, R.B., Wang, Y., Bi, J., Mendes de Moura, Y., Sellers, P.J., 2014. Vegetation dynamics and rainfall sensitivity of the Amazon. Proceedings of the National Academy of Sciences 111, 16041–16046. https://doi.org/10.1073/pnas.1404870111

Hodkinson, I.D., 2005. Terrestrial insects along elevation gradients: species and community responses to altitude. Biological Reviews 80, 489–513. https://doi.org/10.1017/S1464793105006767

Hooten, M.B., Hobbs, N.T., 2015. A guide to Bayesian model selection for ecologists. Ecol Monogr 85, 3–28. https://doi.org/10.1890/14-0661.1

Hubau, W., Lewis, S.L., Phillips, O.L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., Daniels, A.K., Ewango, C.E.N., Fauset, S., Mukinzi, J.M., Sheil, D., Sonké, B., Sullivan, M.J.P., Sunderland, T.C.H., Taedoumg, H., Thomas, S.C., White, L.J.T., Abernethy, K.A., Adu-Bredu, S., Amani, C.A., Baker, T.R., Banin, L.F., Baya, F., Begne, S.K., Bennett, A.C., Benedet, F., Bitariho, R., Bocko, Y.E., Boeckx, P., Boundja, P., Brienen, R.J.W., Brncic, T., Chezeaux, E., Chuyong, G.B., Clark, C.J., Collins, M., Comiskey, J.A., Coomes, D.A., Dargie, G.C., de Haulleville, T., Kamdem, M.N.D., Doucet, J.-L., Esquivel-Muelbert, A., Feldpausch, T.R., Fofanah, A., Foli, E.G., Gilpin, M., Gloor, E., Gonmadje, C., Gourlet-Fleury, S., Hall, J.S., Hamilton, A.C., Harris, D.J., Hart, T.B., Hockemba, M.B.N., Hladik, A., Ifo, S.A., Jeffery, K.J., Jucker, T., Yakusu, E.K., Kearsley, E., Kenfack, D., Koch, A., Leal, M.E., Levesley, A., Lindsell, J.A., Lisingo, J., Lopez-Gonzalez, G., Lovett, J.C., Makana, J.-R., Malhi, Y., Marshall, A.R., Martin, J., Martin, E.H., Mbayu, F.M., Medjibe, V.P., Mihindou, V., Mitchard, E.T.A., Moore, S., Munishi, P.K.T., Bengone, N.N., Ojo, L., Ondo, F.E., Peh, K.S.-H., Pickavance, G.C., Poulsen, A.D., Poulsen, J.R., Qie, L., Reitsma, J., Rovero, F., Swaine, M.D., Talbot, J., Taplin, J., Taylor, D.M., Thomas, D.W., Toirambe, B., Mukendi, J.T., Tuagben, D., Umunay, P.M., van der Heijden, G.M.F., Verbeeck, H., Vleminckx, J., Willcock, S., Wöll, H., Woods, J.T., Zemagho, L., 2020. Asynchronous carbon sink saturation in African and Amazonian tropical forests. Nature 579, 80–87. https://doi.org/10.1038/s41586-020-2035-0

Huete, A.R., 2012. Vegetation Indices, Remote Sensing and Forest Monitoring. Geogr Compass 6, 513–532. https://doi.org/10.1111/j.1749-8198.2012.00507.x

Huete, A.R., HuiQing Liu, van Leeuwen, W.J.D., 1997. The use of vegetation indices in forested regions: issues of linearity and saturation, in: IGARSS’97. 1997 IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development. IEEE, pp. 1966–1968. https://doi.org/10.1109/IGARSS.1997.609169

Humphreys, A.M., Govaerts, R., Ficinski, S.Z., Nic Lughadha, E., Vorontsova, M.S., 2019. Global dataset shows geography and life form predict modern plant extinction and rediscovery. Nat Ecol Evol 3, 1043–1047. https://doi.org/10.1038/s41559-019-0906-2

Iheaturu, C.J., Hepner, S., Batchelor, J.L., Agonvonon, G.A., Akinyemi, F.O., Wingate, V.R., Ifejika Speranza, C., 2024. Integrating UAV LiDAR and multispectral data to assess forest status and map disturbance severity in a West African forest patch. Ecol Inform 84, 102876. https://doi.org/10.1016/j.ecoinf.2024.102876

Inoue, Y., Peñuelas, J., Miyata, A., Mano, M., 2008. Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO2 flux measurements in rice. Remote Sens Environ 112, 156–172. https://doi.org/10.1016/J.RSE.2007.04.011

IPBES, 2016. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. https://doi.org/10.5281/ZENODO.6417333

IPCC, (Intergovernmental Panel on Climate Change), 2023. Climate Change 2021 – The Physical Science Basis, Climate Change 2021 – The Physical Science Basis. Cambridge University Press. https://doi.org/10.1017/9781009157896

Irteza, S.M., Nichol, J.E., Shi, W., Abbas, S., 2021. NDVI and Fluorescence Indicators of Seasonal and Structural Changes in a Tropical Forest Succession. Earth Systems and Environment 5, 127–133. https://doi.org/10.1007/s41748-020-00175-5

J van Proosdij, A.S., M Sosef, Marc S, Wieringa, Jan J, Raes S J van Proosdij andrevanproosdij, N.A., M Sosef, M S, Wieringa, J J, 2016. Minimum required number of specimen records to develop accurate species distribution models. Ecography 39, 542–552. https://doi.org/10.1111/ecog.01509

Jarocińska, A., Kopeć, D., Kycko, M., 2024. Comparison of dimensionality reduction methods on hyperspectral images for the identification of heathlands and mires. Sci Rep 14, 27662. https://doi.org/10.1038/s41598-024-79209-1

Jenkins, C.N., Joppa, L., 2009. Expansion of the global terrestrial protected area system. Biol Conserv 142, 2166–2174. https://doi.org/10.1016/j.biocon.2009.04.016

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M., Moorcroft, P., Schaepman, M.E., Schildhauer, M.P., Schneider, F.D., Schrodt, F., Stahl, U., Ustin, S.L., 2016. Monitoring plant functional diversity from space. Nat Plants 2, 16024. https://doi.org/10.1038/nplants.2016.24

Jetz, W., McGeoch, M.A., Guralnick, R., Ferrier, S., Beck, J., Costello, M.J., Fernandez, M., Geller, G.N., Keil, P., Merow, C., Meyer, C., Muller-Karger, F.E., Pereira, H.M., Regan, E.C., Schmeller, D.S., Turak, E., 2019. Essential biodiversity variables for mapping and monitoring species populations. Nat Ecol Evol 3, 539–551. https://doi.org/10.1038/s41559-019-0826-1

Jetz, W., McPherson, J.M., Guralnick, R.P., 2012. Integrating biodiversity distribution knowledge: toward a global map of life. Trends Ecol Evol 27, 151–159. https://doi.org/10.1016/j.tree.2011.09.007

Justice, C.O., Townshend, J.R.G., Vermote, E.F., Masuoka, E., Wolfe, R.E., Saleous, N., Roy, D.P., Morisette, J.T., 2002. An overview of MODIS Land data processing and product status. Remote Sens Environ 83, 3–15. https://doi.org/10.1016/S0034-4257(02)00084-6

Kery, M., 2008. ESTIMATING ABUNDANCE FROM BIRD COUNTS: BINOMIAL MIXTURE MODELS UNCOVER COMPLEX COVARIATE RELATIONSHIPS. Auk 125, 336–345. https://doi.org/10.1525/auk.2008.06185

Kéry, M., Royle, J.A., 2016. Chapter 6 - Modeling Abundance with Counts of Unmarked Individuals in Closed Populations: Binomial N-mixture Models, in: Kéry, M., Royle, J.A. (Eds.), Applied Hierarchical Modeling in Ecology. Academic Press, Boston, pp. 219–312. https://doi.org/https://doi.org/10.1016/B978-0-12-801378-6.00006-0

Kindt, R., 2020. WorldFlora: An R package for exact and fuzzy matching of plant names against the World Flora Online taxonomic backbone data. Appl Plant Sci 8. https://doi.org/10.1002/aps3.11388

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J., Zimmermann, N.E., O'Hara, R.B., 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. J Biogeogr 39, 2163–2178. https://doi.org/10.1111/j.1365-2699.2011.02663.x

Kissling, W.D., Hardisty, A., García, E.A., Santamaria, M., De Leo, F., Pesole, G., Freyhof, J., Manset, D., Wissel, S., Konijn, J., Los, W., 2015. Towards global interoperability for supporting biodiversity research on essential biodiversity variables (EBVs). Biodiversity 16, 99–107. https://doi.org/10.1080/14888386.2015.1068709

Krijthe, J.H., 2015. Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation.

Kumari, N., Saco, P.M., Rodriguez, J.F., Johnstone, S.A., Srivastava, A., Chun, K.P., Yetemen, O., 2020. The Grass Is Not Always Greener on the Other Side: Seasonal Reversal of Vegetation Greenness in Aspect-Driven Semiarid Ecosystems. Geophys Res Lett 47, e2020GL088918. https://doi.org/10.1029/2020GL088918

Kutner, M.H., 2005. Applied linear statistical models, 5th ed. McGraw-Hill.

Langhammer, P.F., Bull, J.W., Bicknell, J.E., Oakley, J.L., Brown, M.H., Bruford, M.W., Butchart, S.H.M., Carr, J.A., Church, D., Cooney, R., Cutajar, S., Foden, W., Foster, M.N., Gascon, C., Geldmann, J., Genovesi, P., Hoffmann, M., Howard-McCombe, J., Lewis, T., Macfarlane, N.B.W., Melvin, Z.E., Merizalde, R.S., Morehouse, M.G., Pagad, S., Polidoro, B., Sechrest, W., Segelbacher, G., Smith, K.G., Steadman, J., Strongin, K., Williams, J., Woodley, S., Brooks, T.M., 2024. The positive impact of conservation action. Science (1979) 384, 453–458. https://doi.org/10.1126/science.adj6598

Lapola, D.M., Pinho, P., Barlow, J., Aragão, L.E.O.C., Berenguer, E., Carmenta, R., Liddy, H.M., Seixas, H., Silva, C.V.J., Silva-Junior, C.H.L., Alencar, A.A.C., Anderson, L.O., Armenteras, D., Brovkin, V., Calders, K., Chambers, J., Chini, L., Costa, M.H., Faria, B.L., Fearnside, P.M., Ferreira, J., Gatti, L., Gutierrez-Velez, V.H., Han, Z., Hibbard, K., Koven, C., Lawrence, P., Pongratz, J., Portela, B.T.T., Rounsevell, M., Ruane, A.C., Schaldach, R., da Silva, S.S., von Randow, C., Walker, W.S., 2023. The drivers and impacts of Amazon forest degradation. Science (1979) 379. https://doi.org/10.1126/science.abp8622

Lasko, K., 2022. Gap Filling Cloudy Sentinel-2 NDVI and NDWI Pixels with Multi-Frequency Denoised C-Band and L-Band Synthetic Aperture Radar (SAR), Texture, and Shallow Learning Techniques. Remote Sens (Basel) 14, 4221. https://doi.org/10.3390/rs14174221

Laurance, W.F., Albernaz, A.K.M., Costa, C. Da, 2001. Is deforestation accelerating in the Brazilian Amazon? Environ Conserv 28, 305–311. https://doi.org/10.1017/S0376892901000339

Laurance, W.F., Camargo, J.L.C., Luizão, R.C.C., Laurance, S.G., Pimm, S.L., Bruna, E.M., Stouffer, P.C., Bruce Williamson, G., Benítez-Malvido, J., Vasconcelos, H.L., Van Houtan, K.S., Zartman, C.E., Boyle, S.A., Didham, R.K., Andrade, A., Lovejoy, T.E., 2011. The fate of Amazonian forest fragments: A 32-year investigation. Biol Conserv 144, 56–67. https://doi.org/10.1016/j.biocon.2010.09.021

Laurance, W.F., Carolina Useche, D., Rendeiro, J., Kalka, M., Bradshaw, C.J.A., 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., Cords, M., Danielsen, F., De Dijn, B., Dinerstein, E., Donnelly, M.A., Edwards, D., Edwards, F., Farwig, N., Fashing, P., Forget, P.-M., Foster, M., Gale, G., Harris, D., Harrison, R., Hart, J., Karpanty, S., John Kress, W., Krishnaswamy, J., Logsdon, W., Lovett, J., Magnusson, W., Maisels, F., Marshall, A.R., McClearn, D., Mudappa, D., Nielsen, M.R., Pearson, R., Pitman, N., van der Ploeg, J., Plumptre, A., Poulsen, J., Quesada, M., Rainey, H., Robinson, D., Roetgers, C., Rovero, F., Scatena, F., Schulze, C., Sheil, D., Struhsaker, T., Terborgh, J., Thomas, D., Timm, R., Nicolas Urbina-Cardona, J., Vasudevan, K., Joseph Wright, S., Carlos Arias-G., J., Arroyo, L., Ashton, M., Auzel, P., Babaasa, D., Babweteera, F., Baker, P., Banki, O., Bass, M., Bila-Isia, I., Blake, S., Brockelman, W., Brokaw, N., Brühl, C.A., Bunyavejchewin, S., Chao, J.-T., Chave, J., Chellam, R., Clark, C.J., Clavijo, J., Congdon, R., Corlett, R., Dattaraja, H.S., Dave, C., Davies, G., de Mello Beisiegel, B., de Nazaré Paes da Silva, R., Di Fiore, A., Diesmos, A., Dirzo, R., Doran-Sheehy, D., Eaton, M., Emmons, L., Estrada, A., Ewango, C., Fedigan, L., Feer, F., Fruth, B., Giacalone Willis, J., Goodale, U., Goodman, S., Guix, J.C., Guthiga, P., Haber, W., Hamer, K., Herbinger, I., Hill, J., Huang, Z., Fang Sun, I., Ickes, K., Itoh, A., Ivanauskas, N., Jackes, B., Janovec, J., Janzen, D., Jiangming, M., Jin, C., Jones, T., Justiniano, H., Kalko, E., Kasangaki, A., Killeen, T., King, H., Klop, E., Knott, C., Koné, I., Kudavidanage, E., Lahoz da Silva Ribeiro, J., Lattke, J., Laval, R., Lawton, R., Leal, M., Leighton, M., Lentino, M., Leonel, C., Lindsell, J., Ling-Ling, L., Eduard Linsenmair, K., Losos, E., Lugo, A., Lwanga, J., Mack, A.L., Martins, M., Scott McGraw, W., McNab, R., Montag, L., Myers Thompson, J., Nabe-Nielsen, J., Nakagawa, M., Nepal, S., Norconk, M., Novotny, V., O'Donnell, S., Opiang, M., Ouboter, P., Parker, K., Parthasarathy, N., Pisciotta, K., Prawiradilaga, D., Pringle, C., Rajathurai, S., Reichard, U., Reinartz, G., Renton, K., Reynolds, G., Reynolds, V., Riley, E., Rödel, M.-O., Rothman, J., Round, P., Sakai, S., Sanaiotti, T., Savini, T., Schaab, G., Seidensticker, J., Siaka, A., Silman, M.R., Smith, T.B., de Almeida, S.S., Sodhi, N., Stanford, C., Stewart, K., Stokes, E., Stoner, K.E., Sukumar, R., Surbeck, M., Tobler, M., Tscharntke, T., Turkalo, A., Umapathy, G., van Weerd, M., Vega Rivera, J., Venkataraman, M., Venn, L., Verea, C., Volkmer de Castilho, C., Waltert, M., Wang, B., Watts, D., Weber, W., West, P., Whitacre, D., Whitney, K., Wilkie, D., Williams, S., Wright, D.D., Wright, P., Xiankai, L., Yonzon, P., Zamzani, F., 2012. Averting biodiversity collapse in tropical forest protected areas. Nature 489, 290–294. https://doi.org/10.1038/nature11318

Laurance, W.F., Sayer, J., Cassman, K.G., 2014. Agricultural expansion and its impacts on tropical nature. Trends Ecol Evol 29, 107–116. https://doi.org/10.1016/j.tree.2013.12.001

Lausch, A., Bannehr, L., Beckmann, M., Boehm, C., Feilhauer, H., Hacker, J.M., Heurich, M., Jung, A., Klenke, R., Neumann, C., Pause, M., Rocchini, D., Schaepman, M.E., Schmidtlein, S., Schulz, K., Selsam, P., Settele, J., Skidmore, A.K., Cord, A.F., 2016. Linking Earth Observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. Ecol Indic 70, 317–339. https://doi.org/10.1016/j.ecolind.2016.06.022

Lausch, A., Blaschke, T., Haase, D., Herzog, F., Syrbe, R.-U., Tischendorf, L., Walz, U., 2015. Understanding and quantifying landscape structure – A review on relevant process characteristics, data models and landscape metrics. Ecol Modell 295, 31–41. https://doi.org/10.1016/j.ecolmodel.2014.08.018

Lausch, A., Pause, M., Schmidt, A., Salbach, C., Gwillym-Margianto, S., Merbach, I., 2013. Temporal hyperspectral monitoring of chlorophyll, LAI, and water content of barley during a growing season. Canadian Journal of Remote Sensing 39, 191–207. https://doi.org/10.5589/m13-028

Le Provost, G., Schenk, N. V., Penone, C., Thiele, J., Westphal, C., Allan, E., Ayasse, M., Blüthgen, N., Boeddinghaus, R.S., Boesing, A.L., Bolliger, R., Busch, V., Fischer, M., Gossner, M.M., Hölzel, N., Jung, K., Kandeler, E., Klaus, V.H., Kleinebecker, T., Leimer, S., Marhan, S., Morris, K., Müller, S., Neff, F., Neyret, M., Oelmann, Y., Perović, D.J., Peter, S., Prati, D., Rillig, M.C., Saiz, H., Schäfer, D., Scherer-Lorenzen, M., Schloter, M., Schöning, I., Schrumpf, M., Steckel, J., Steffan-Dewenter, I., Tschapka, M., Vogt, J., Weiner, C., Weisser, W., Wells, K., Werner, M., Wilcke, W., Manning, P., 2022. The supply of multiple ecosystem services requires biodiversity across spatial scales. Nat Ecol Evol. https://doi.org/10.1038/s41559-022-01918-5

Leibold, M.A., Rudolph, F.J., Blanchet, F.G., De Meester, L., Gravel, D., Hartig, F., Peres-Neto, P., Shoemaker, L., Chase, J.M., 2022. The internal structure of metacommunities. Oikos 2022. https://doi.org/10.1111/oik.08618

Leitão, P.J., Schwieder, M., Senf, C., 2017. Sgdm: An R package for performing sparse generalized dissimilarity modelling with tools for gdm. ISPRS Int J Geoinf 6, 23. https://doi.org/10.3390/ijgi6010023

Leitão, P.J., Schwieder, M., Senf, C., 2016. sgdm: Package for performing sparse generalized dissimilarity model including tools for gdm.

Leitão, P.J., Schwieder, M., Suess, S., Catry, I., Milton, E.J., Moreira, F., Osborne, P.E., Pinto, M.J., Linden, S., Hostert, P., 2015. Mapping beta diversity from space: Sparse Generalised Dissimilarity Modelling (SGDM) for analysing high-dimensional data. Methods Ecol Evol 6, 764–771. https://doi.org/10.1111/2041-210X.12378

Lennox, G.D., Gardner, T.A., Thomson, J.R., Ferreira, J., Berenguer, E., Lees, A.C., Mac Nally, R., Aragão, L.E.O.C., Ferraz, S.F.B., Louzada, J., Moura, N.G., Oliveira, V.H.F., Pardini, R., Solar, R.R.C., Vaz-de Mello, F.Z., Vieira, I.C.G., Barlow, J., 2018. Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. Glob Chang Biol 24, 5680–5694. https://doi.org/10.1111/gcb.14443

Lennox, R.J., Suski, C.D., Cooke, S.J., 2018. A macrophysiology approach to watershed science and management. Science of The Total Environment 626, 434–440. https://doi.org/10.1016/j.scitotenv.2018.01.069

Lepage, D., 2024. Tambopata bird checklist - Avibase - Bird Checklists of the World [WWW Document]. URL https://avibase.bsc-eoc.org/checklist.jsp?lang=EN&p2=1&list=howardmoore&synlang=&region=PEmdtp&version=text&lifelist=&highlight=0 (accessed 2.10.25).

Lewis, S.L., Edwards, D.P., Galbraith, D., 2015. Increasing human dominance of tropical forests. Science (1979) 349, 827–832. https://doi.org/10.1126/science.aaa9932

Li, Y., Devenish, C., Tosa, M.I., Luo, M., Bell, D.M., Lesmeister, D.B., Greenfield, P., Pichler, M., Levi, T., Yu, D.W., 2024. Combining environmental DNA and remote sensing for efficient, fine-scale mapping of arthropod biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences 379. https://doi.org/10.1098/rstb.2023.0123

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Lopes, M., Frison, P., Durant, S.M., Schulte to Bühne, H., Ipavec, A., Lapeyre, V., Pettorelli, N., 2020. Combining optical and radar satellite image time series to map natural vegetation: savannas as an example. Remote Sens Ecol Conserv 6, 316–326. https://doi.org/10.1002/rse2.139

Ludwig, A., Meyer, H., Nauss, T., 2016. Automatic classification of Google Earth images for a larger scale monitoring of bush encroachment in South Africa. International Journal of Applied Earth Observation and Geoinformation 50, 89–94. https://doi.org/10.1016/j.jag.2016.03.003

Mace, G.M., Norris, K., Fitter, A.H., 2012. Biodiversity and ecosystem services: a multilayered relationship. Trends Ecol Evol 27, 19–26. https://doi.org/10.1016/j.tree.2011.08.006

MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., Franklin, A.B., 2003. ESTIMATING SITE OCCUPANCY, COLONIZATION, AND LOCAL EXTINCTION WHEN A SPECIES IS DETECTED IMPERFECTLY. Ecology 84, 2200–2207. https://doi.org/10.1890/02-3090

Magioli, M., Ferraz, K.M.P.M. de B., Chiarello, A.G., Galetti, M., Setz, E.Z.F., Paglia, A.P., Abrego, N., Ribeiro, M.C., Ovaskainen, O., 2021. Land-use changes lead to functional loss of terrestrial mammals in a Neotropical rainforest. Perspect Ecol Conserv 19, 161–170. https://doi.org/10.1016/j.pecon.2021.02.006

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.McP., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P.J., Watt, A.D., 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. Trends Ecol Evol 25, 574–582. https://doi.org/10.1016/j.tree.2010.06.016

Malhi, Y., Aragão, L.E.O.C., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., Sitch, S., McSweeney, C., Meir, P., 2009. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. Proceedings of the National Academy of Sciences 106, 20610–20615. https://doi.org/10.1073/pnas.0804619106

Maliniemi, T., Happonen, K., Virtanen, R., 2019. Site fertility drives temporal turnover of vegetation at high latitudes. Ecol Evol 9, 13255–13266. https://doi.org/10.1002/ece3.5778

Mandrekar, J.N., 2010. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. Journal of Thoracic Oncology 5, 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

MapBiomas, 2023. MapBiomas Peru [WWW Document]. https://peru.mapbiomas.org/en/terminos-de-uso/. URL https://peru.mapbiomas.org/en/colecciones-de-mapbiomas-peru/ (accessed 8.21.23).

Marcus, M., Gutierrez-Velez, V.H., Cronkleton, P., 2020. Land use change in four landscapes in the Peruvian Amazon, Land use change in four landscapes in the Peruvian Amazon. Center for International Forestry Research (CIFOR). https://doi.org/10.17528/cifor/007736

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

Martin, T.E., Nightingale, J., Baddams, J., Monkhouse, J., Kaban, A., Sastranegara, H., Mulyani, Y., Blackburn, G.A., Simcox, W., 2017. Variability in the Effectiveness of Two Ornithological Survey Methods between Tropical Forest Ecosystems. PLoS One 12, e0169786. https://doi.org/10.1371/journal.pone.0169786

Matos, F.A.R., Magnago, L.F.S., Aquila Chan Miranda, C., de Menezes, L.F.T., Gastauer, M., Safar, N.V.H., Schaefer, C.E.G.R., da Silva, M.P., Simonelli, M., Edwards, F.A., Martins, S. V., Meira-Neto, J.A.A., Edwards, D.P., 2020. Secondary forest fragments offer important carbon and biodiversity cobenefits. Glob Chang Biol 26, 509–522. https://doi.org/10.1111/gcb.14824

Maxwell, S.L., Fuller, R.A., Brooks, T.M., Watson, J.E.M., 2016. Biodiversity: The ravages of guns, nets and bulldozers. Nature 536, 143–145. https://doi.org/10.1038/536143a

McCord, S.E., Webb, N.P., Van Zee, J.W., Burnett, S.H., Christensen, E.M., Courtright, E.M., Laney, C.M., Lunch, C., Maxwell, C., Karl, J.W., Slaughter, A., Stauffer, N.G., Tweedie, C., 2021. Provoking a Cultural Shift in Data Quality. Bioscience 71, 647–657. https://doi.org/10.1093/biosci/biab020

McCoy, E.D., 1990. The Distribution of Insects along Elevational Gradients. Oikos 58, 313. https://doi.org/10.2307/3545222

McDonald, J.A., Helmstedt, K.J., Bode, M., Coutts, S., McDonald-Madden, E., Possingham, H.P., 2018. Improving private land conservation with outcome-based biodiversity payments. Journal of Applied Ecology 55, 1476–1485. https://doi.org/10.1111/1365-2664.13071

McDowell, N.G., Coops, N.C., Beck, P.S.A., Chambers, J.Q., Gangodagamage, C., Hicke, J.A., Huang, C. ying, Kennedy, R., Krofcheck, D.J., Litvak, M., Meddens, A.J.H., Muss, J., Negrón-Juarez, R., Peng, C., Schwantes, A.M., Swenson, J.J., Vernon, L.J., Williams, A.P., Xu, C., Zhao, M., Running, S.W., Allen, C.D., 2015. Global satellite monitoring of climate-induced vegetation disturbances. Trends Plant Sci. https://doi.org/10.1016/j.tplants.2014.10.008

McElreath, R., 2020. Statistical Rethinking. Chapman and Hall/CRC. https://doi.org/10.1201/9780429029608

Mcfeeters, S.K., 1999. Remote Sensing Using the Normalized Difference Water Index (NDWI) within a Geographic Information System to Detect Swimming Pools for Mosquito Abatement: A Practical Approach. Remote Sens 5, 3544–3561. https://doi.org/10.3390/rs5073544

McGarigal, K., Wan, H.Y., Zeller, K.A., Timm, B.C., Cushman, S.A., 2016. Multi-scale habitat selection modeling: a review and outlook. Landsc Ecol 31, 1161–1175. https://doi.org/10.1007/s10980-016-0374-x

Metcalf, O.C., Barlow, J., Devenish, C., Marsden, S., Berenguer, E., Lees, A.C., 2021. Acoustic indices perform better when applied at ecologically meaningful time and frequency scales. Methods Ecol Evol 12, 421–431. https://doi.org/10.1111/2041-210X.13521

Metcalf, O.C., Barlow, J., Marsden, S., Gomes de Moura, N., Berenguer, E., Ferreira, J., Lees, A.C., 2022. Optimizing tropical forest bird surveys using passive acoustic monitoring and high temporal resolution sampling. Remote Sens Ecol Conserv 8, 45–56. https://doi.org/10.1002/rse2.227

Mishra, V.N., Prasad, R., Rai, P.K., Vishwakarma, A.K., Arora, A., 2019. Performance evaluation of textural features in improving land use/land cover classification accuracy of heterogeneous landscape using multi-sensor remote sensing data. Earth Sci Inform 12, 71–86. https://doi.org/10.1007/S12145-018-0369-Z/TABLES/10

Mitchard, E.T.A., Saatchi, S.S., Lewis, S.L., Feldpausch, T.R., Woodhouse, I.H., Sonké, B., Rowland, C., Meir, P., 2011. Measuring biomass changes due to woody encroachment and deforestation/degradation in a forest–savanna boundary region of central Africa using multi-temporal L-band radar backscatter. Remote Sens Environ 115, 2861–2873. https://doi.org/10.1016/j.rse.2010.02.022

Mitchell, R.M., Bakker, J.D., Vincent, J.B., Davies, G.M., 2017. Relative importance of abiotic, biotic, and disturbance drivers of plant community structure in the sagebrush steppe. Ecological Applications 27, 756–768. https://doi.org/10.1002/eap.1479

Mod, H.K., Scherrer, D., Luoto, M., Guisan, A., 2016. What we use is not what we know: environmental predictors in plant distribution models. Journal of Vegetation Science 27, 1308–1322. https://doi.org/10.1111/JVS.12444

Morrison, L.W., 2016. Observer error in vegetation surveys: a review. Journal of Plant Ecology 9, 367–379. https://doi.org/10.1093/jpe/rtv077

Morton, D.C., DeFries, R.S., Shimabukuro, Y.E., Anderson, L.O., Del Bon Espírito-Santo, F., Hansen, M., Carroll, M., 2005. Rapid Assessment of Annual Deforestation in the Brazilian Amazon Using MODIS Data. Earth Interact 9, 1–22. https://doi.org/10.1175/EI139.1

Moudrý, V., Šímová, P., 2012. Influence of positional accuracy, sample size and scale on modelling species distributions: a review. International Journal of Geographical Information Science 26, 2083–2095. https://doi.org/10.1080/13658816.2012.721553

Mulvaney, J.M., Cherry, M.I., 2020. The effectiveness of point counts and mist-netting in surveying Afrotemperate forest bird community structure in South Africa. Emu - Austral Ornithology 120, 103–113. https://doi.org/10.1080/01584197.2020.1726186

Musker, S.D., Ellis, A.G., Schlebusch, S.A., Verboom, G.A., 2021. Niche specificity influences gene flow across fine-scale habitat mosaics in Succulent Karoo plants. Mol Ecol 30, 175–192. https://doi.org/10.1111/mec.15721

Myers, N., Mittermeler, R.A., Mittermeler, C.G., Da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858. https://doi.org/10.1038/35002501

Nawrocki, T.W., Carlson, M.L., Osnas, J.L.D., Trammell, E.J., Witmer, F.D.W., 2020. Regional mapping of species-level continuous foliar cover: beyond categorical vegetation mapping. Ecological Applications 30. https://doi.org/10.1002/eap.2081

Nazarova, T., Martin, P., Giuliani, G., 2020. Monitoring Vegetation Change in the Presence of High Cloud Cover with Sentinel-2 in a Lowland Tropical Forest Region in Brazil. Remote Sens (Basel) 12, 1829. https://doi.org/10.3390/rs12111829

Newbold, T., Hudson, L.N., Arnell, A.P., Contu, S., De Palma, A., Ferrier, S., Hill, S.L.L., Hoskins, A.J., Lysenko, I., Phillips, H.R.P., Burton, V.J., Chng, C.W.T., Emerson, S., Gao, D., Pask-Hale, G., Hutton, J., Jung, M., Sanchez-Ortiz, K., Simmons, B.I., Whitmee, S., Zhang, H., Scharlemann, J.P.W., Purvis, A., 2016. Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. Science (1979) 353, 288–291. https://doi.org/10.1126/science.aaf2201

Nguyen Trong, H., Nguyen, T.D., Kappas, M., 2020. Land Cover and Forest Type Classification by Values of Vegetation Indices and Forest Structure of Tropical Lowland Forests in Central Vietnam. International Journal of Forestry Research 2020, 1–18. https://doi.org/10.1155/2020/8896310

Nichols, J., Williams, B., 2006. Monitoring for conservation. Trends Ecol Evol 21, 668–673. https://doi.org/10.1016/j.tree.2006.08.007

Niemi, G.J., McDonald, M.E., 2004. Application of Ecological Indicators. Annu Rev Ecol Evol Syst 35, 89–111. https://doi.org/10.1146/annurev.ecolsys.35.112202.130132

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., Foster, S.D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N.A., Holt, R.D., Hui, F.K.C., Husby, M., Kålås, J.A., Lehikoinen, A., Luoto, M., Mod, H.K., Newell, G., Renner, I., Roslin, T., Soininen, J., Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N.E., Gravel, D., Ovaskainen, O., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecol Monogr 89, e01370. https://doi.org/10.1002/ecm.1370

Nores, M., 2000. Species Richness in the Amazonian Bird Fauna from an Evolutionary Perspective. Emu - Austral Ornithology 100, 419–430. https://doi.org/10.1071/MU0007S

Noss, R.F., 1990. Indicators for Monitoring Biodiversity: A Hierarchical Approach. Conservation Biology 4, 355–364. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x

Numbisi, F.N., Van Coillie, F.M.B., De Wulf, R., 2019. Delineation of Cocoa Agroforests Using Multiseason Sentinel-1 SAR Images: A Low Grey Level Range Reduces Uncertainties in GLCM Texture-Based Mapping. ISPRS Int J Geoinf 8, 179. https://doi.org/10.3390/ijgi8040179

Oliver, T.H., Heard, M.S., Isaac, N.J.B., Roy, D.B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A., Orme, C.D.L., Petchey, O.L., Proença, V., Raffaelli, D., Suttle, K.B., Mace, G.M., Martín-López, B., Woodcock, B.A., Bullock, J.M., 2015. Biodiversity and Resilience of Ecosystem Functions. Trends Ecol Evol 30, 673–684. https://doi.org/10.1016/j.tree.2015.08.009

Ollinger, S. V., 2011. Sources of variability in canopy reflectance and the convergent properties of plants. New Phytologist 189, 375–394. https://doi.org/10.1111/j.1469-8137.2010.03536.x

Olofsson, P., Foody, G.M., Herold, M., Stehman, S. V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens Environ 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Ovaskainen, O., Abrego, N., 2020. Joint Species Distribution Modelling, in: Joint Species Distribution Modelling. Cambridge University Press, pp. 142–183. https://doi.org/10.1017/9781108591720.010

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., Abrego, N., 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecol Lett 20, 561–576. https://doi.org/10.1111/ele.12757

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. Science (1979) 333, 988–993. https://doi.org/10.1126/SCIENCE.1201609/SUPPL_FILE/PAPV2.PDF

Parikh, G., Rawtani, D., Khatri, N., 2021. "Insects as an Indicator for Environmental Pollution." Environmental Claims Journal 33, 161–181. https://doi.org/10.1080/10406026.2020.1780698

Parolin, P., 2006. Ombrohydrochory: Rain-operated seed dispersal in plants – With special regard to jet-action dispersal in Aizoaceae. Flora - Morphology, Distribution, Functional Ecology of Plants 201, 511–518. https://doi.org/10.1016/j.flora.2005.11.003

Peh, K.S. -H., Sodhi, N.S., De Jong, J., Sekercioglu, C.H., Yap, C.A. -M., Lim, S.L. -H., 2006. Conservation value of degraded habitats for forest birds in southern Peninsular Malaysia. Divers Distrib 12, 572–581. https://doi.org/10.1111/j.1366-9516.2006.00257.x

Peña-Lara, V.A., Dupuy, J.M., Reyes-Garcia, C., Sanaphre-Villanueva, L., Portillo-Quintero, C.A., Hernández-Stefanoni, J.L., 2022. Modelling Species Richness and Functional Diversity in Tropical Dry Forests Using Multispectral Remotely Sensed and Topographic Data. Remote Sens (Basel) 14, 5919. https://doi.org/10.3390/rs14235919

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M., Wegmann, M., 2013. Essential Biodiversity Variables. Science (1979) 339, 277–278. https://doi.org/10.1126/science.1229931

Pereira Mendes, C., Lim, N.T.-L., 2024. EcoLiDAR: An economical LiDAR scanner for ecological research. PLoS One 19, e0298712. https://doi.org/10.1371/journal.pone.0298712

Perret, J., Besnard, A., Charpentier, A., Papuga, G., 2023. Plants stand still but hide: Imperfect and heterogeneous detection is the rule when counting plants. Journal of Ecology 00, 1–14. https://doi.org/10.1111/1365-2745.14110

Perrone, M., Di Febbraro, M., Conti, L., Divíšek, J., Chytrý, M., Keil, P., Carranza, M.L., Rocchini, D., Torresani, M., Moudrý, V., Šímová, P., Prajzlerová, D., Müllerová, J., Wild, J., Malavasi, M., 2023. The relationship between spectral and plant diversity:

Disentangling the influence of metrics and habitat types at the landscape scale. Remote Sens Environ 293, 113591. https://doi.org/10.1016/j.rse.2023.113591

Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H., Turner, W., Milner-Gulland, E.J., 2014a. Satellite remote sensing for applied ecologists: opportunities and challenges. Journal of Applied Ecology 51, 839–848. https://doi.org/10.1111/1365-2664.12261

Pettorelli, N., Safi, K., Turner, W., 2014b. Satellite remote sensing, biodiversity research and conservation of the future. Philosophical Transactions of the Royal Society B: Biological Sciences 369, 20130190. https://doi.org/10.1098/rstb.2013.0190

Pettorelli, N., Schulte to Bühne, H., 2022. Current and future opportunities for satellite remote sensing to inform rewilding. Remote Sens Ecol Conserv. https://doi.org/10.1002/rse2.321

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.Chr., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol Evol 20, 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Pfeifer, M., Lefebvre, V., Peres, C.A., Banks-Leite, C., Wearn, O.R., Marsh, C.J., Butchart, S.H.M., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley, A., Harris, S.M., Klingbeil, B.T., Kormann, U., Lens, L., Medina-Rangel, G.F., Morante-Filho, J.C., Olivier, P., Peters, S.L., Pidgeon, A., Ribeiro, D.B., Scherber, C., Schneider-Maunoury, L., Struebig, M., Urbina-Cardona, N., Watling, J.I., Willig, M.R., Wood, E.M., Ewers, R.M., 2017. Creation of forest edges has a global impact on forest vertebrates. Nature 551, 187–191. https://doi.org/10.1038/nature24457

Phalan, B., Onial, M., Balmford, A., Green, R.E., 2011. Reconciling Food Production and Biodiversity Conservation: Land Sharing and Land Sparing Compared. Science (1979) 333, 1289–1291. https://doi.org/10.1126/science.1208742

Pichler, M., Hartig, F., 2021. A new joint species distribution model for faster and more accurate inference of species associations from big community data. Methods Ecol Evol 12, 2159–2173. https://doi.org/10.1111/2041-210X.13687

Pimm, S., Raven, P., Peterson, A., Şekercioğlu, Ç.H., Ehrlich, P.R., 2006. Human impacts on the rates of recent, present, and future bird extinctions. Proc Natl Acad Sci U S A 103, 10941–10946. https://doi.org/10.1073/PNAS.0604181103/

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M., Sexton, J.O., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science (1979) 344. https://doi.org/10.1126/science.1246752

Pinto, S., Benincà, E., van Nes, E.H., Scheffer, M., Bogaards, J.A., 2022. Species abundance correlations carry limited information about microbial network interactions. PLoS Comput Biol 18, e1010491. https://doi.org/10.1371/journal.pcbi.1010491

Pinto-Ledezma, J.N., Cavender-Bares, J., 2021. Predicting species distributions and community composition using satellite remote sensing predictors. Sci Rep 11, 16448. https://doi.org/10.1038/s41598-021-96047-7

Piper, A.M., Batovska, J., Cogan, N.O.I., Weiss, J., Cunningham, J.P., Rodoni, B.C., Blacket, M.J., 2019. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. Gigascience 8, 1–22. https://doi.org/10.1093/gigascience/giz092

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., Pélissier, R., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nat Commun 11, 4540. https://doi.org/10.1038/s41467-020-18321-y

Podani, J., 2006. Braun-Blanquet's legacy and data analysis in vegetation science. Journal of Vegetation Science 17, 113–117. https://doi.org/10.1111/j.1654-1103.2006.tb02429.x

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A., McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model. Methods Ecol Evol 5, 397–406. https://doi.org/10.1111/2041-210X.12180

Potapov, P., Hansen, M.C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., Esipova, E., 2017. The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. Sci Adv 3. https://doi.org/10.1126/SCIADV.1600821/ASSET/65C1F2AE-AAD4-48F4-928F-9CA0C57BAE72/ASSETS/GRAPHIC/1600821-F7.JPEG

Purdon, A., Mole, M.A., Selier, J., Kruger, J., Mafumo, H., Olivier, P.I., 2022. Using the Rao's Q diversity index as an indicator of protected area effectiveness in conserving biodiversity. Ecol Inform 72, 101920. https://doi.org/10.1016/j.ecoinf.2022.101920

Pureswaran, D.S., Roques, A., Battisti, A., 2018. Forest Insects and Climate Change. Current Forestry Reports 4, 35–50. https://doi.org/10.1007/s40725-018-0075-6

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

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.-F., Ho Tong Minh, D., Lomas, M., Mariotti, M.D., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Luke Smallman, T., Soja, M.J., Tebaldini, S., Ulander, L., Villard, L., Williams, M., 2019. The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. Remote Sens Environ 227, 44–60. https://doi.org/10.1016/j.rse.2019.03.032

Quiñones, M., Vissers, M., Hoekman, D., 2007. K&C Science Report - Phase 1 Tropical Forest and Wetlands Mapping, Case Study Borneo.

R Core Team, 2022. R: A Language and Environment for Statistical Computing.

Rappaport, D.I., Morton, D.C., Longo, M., Keller, M., Dubayah, R., dos-Santos, M.N., 2018. Quantifying long-term changes in carbon stocks and forest structure from Amazon forest degradation. Environmental Research Letters 13, 065013. https://doi.org/10.1088/1748-9326/aac331

Rappaport, D.I., Swain, A., Fagan, W.F., Dubayah, R., Morton, D.C., 2022. Animal soundscapes reveal key markers of Amazon forest degradation from fire and logging. Proc Natl Acad Sci U S A 119, 805–4248. https://doi.org/10.1073/pnas.2102878119

Regos, A., Gómez-Rodríguez, P., Arenas-Castro, S., Tapia, L., Vidal, M., Domínguez, J., 2020. Model-Assisted Bird Monitoring Based on Remotely Sensed Ecosystem Functioning and Atlas Data. Remote Sens (Basel) 12, 2549. https://doi.org/10.3390/rs12162549

Regos, A., Gonçalves, J., Arenas-Castro, S., Alcaraz-Segura, D., Guisan, A., Honrado, J.P., 2022. Mainstreaming remotely sensed ecosystem functioning in ecological niche models. Remote Sens Ecol Conserv 8, 431–447. https://doi.org/10.1002/rse2.255

Remsen, J. V., Good, D.A., 1996. Misuse of Data from Mist-Net Captures to Assess Relative Abundance in Bird Populations. Auk 113, 381–398. https://doi.org/10.2307/4088905

Robinson, W.D., Curtis, J.R., 2020. Creating benchmark measurements of tropical forest bird communities in large plots. Condor 122, 1–15. https://doi.org/10.1093/condor/duaa015

Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M., Gillespie, T.W., He, K.S., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland, J., Ricotta, C., Southworth, J., Neteler, M., 2010. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. Ecol Inform 5, 318–329. https://doi.org/10.1016/j.ecoinf.2010.06.001

Rocchini, D., Boyd, D.S., Féret, J., Foody, G.M., He, K.S., Lausch, A., Nagendra, H., Wegmann, M., Pettorelli, N., 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. Remote Sens Ecol Conserv 2, 25–36. https://doi.org/10.1002/rse2.9

Rocchini, D., Chiarucci, A., Loiselle, S.A., 2004. Testing the spectral variation hypothesis by using satellite multispectral images. Acta Oecologica 26, 117–120. https://doi.org/10.1016/j.actao.2004.03.008

Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., Feilhauer, H., Féret, J.B., Foody, G.M., Gavish, Y., Godinho, S., Kunin, W.E., Lausch, A., Leitão, P.J., Marcantonio, M., Neteler, M., Ricotta, C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H., 2018. Measuring β-diversity by remote sensing: A challenge for biodiversity monitoring. Methods Ecol Evol 9, 1787–1798. https://doi.org/10.1111/2041-210X.12941

Rosenstock, S.S., Anderson, D.R., Giesen, K.M., Leukering, T., Carter, M.F., 2002. Landbird Counting Techniques: Current Practices and an Alternative. Auk 119, 46–53. https://doi.org/10.1093/auk/119.1.46

Ross, S.R.P. -J., Petchey, O.L., Sasaki, T., Armitage, D.W., 2023. How to measure response diversity. Methods Ecol Evol 14, 1150–1167. https://doi.org/10.1111/2041-210X.14087

Rouse, J.W., Hass, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the Great Plains with ERTS, in: Nasa ERTS Symposium. pp. 309–313.

Royle, J.A., Dorazio, R.M., 2009. Hierarchical Modeling and Inference in Ecology, Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities. Elsevier. https://doi.org/10.1016/B978-0-12-374097-7.X0001-4

Runke, W., Xiaoni, Y., Yaya, S., Chengyong, W., Baokang, L., 2022. Study on air temperature estimation and its influencing factors in a complex mountainous area. PLoS One 17, e0272946. https://doi.org/10.1371/journal.pone.0272946

Rutt, C.L., Jirinec, V., Cohn-Haft, M., Laurance, W.F., Stouffer, P.C., 2019. Avian ecological succession in the Amazon: A long-term case study following experimental deforestation. Ecol Evol 9, 13850–13861. https://doi.org/10.1002/ece3.5822

Rutt, C.L., Stouffer, P.C., 2021. Seasonal dynamics of flock interaction networks across a human-modified landscape in lowland Amazonian rain forest. Ecological Applications 31. https://doi.org/10.1002/eap.2235

Saayman, N., Morris, C., Botha, H., 2016. Livestock grazing has minimal effect on the species richness and diversity of semi-arid shrubland in the Western Little Karoo, South Africa. Afr J Range Forage Sci 33, 225–233. https://doi.org/10.2989/10220119.2016.1243581

Samarawickrama, U., Piyaratne, D., Ranagalage, M., 2017. Relationship between NDVI with Tasseled cap Indices: A Remote Sensing based Analysis 3.

SANBI, 2025. CAPE Programme - SANBI [WWW Document]. URL https://www.sanbi.org/biodiversity/science-into-policy-action/mainstreaming-biodiversity/cape-programme/ (accessed 1.20.25).

SANBI, 2019. The Vegetation Map of South Africa, Lesotho and Swaziland (2006-2018). http://bgis.sanbi.org/Projects/Detail/186.

Savilaakso, S., Lausberg, N., Waeber, P.O., Hillgén, O., Isotalo, A., Kleinschroth, F., Djenontin, I.N.S., Boul Lefeuvre, N., Garcia, C.A., 2023. Whose perspective counts? A critical look at definitions of terms used for natural and near-natural forests. One Earth 6, 1477–1493. https://doi.org/10.1016/j.oneear.2023.10.003

Schmeller, D.S., Böhm, M., Arvanitidis, C., Barber-Meyer, S., Brummitt, N., Chandler, M., Chatzinikolaou, E., Costello, M.J., Ding, H., García-Moreno, J., Gill, M., Haase, P., Jones, M., Juillard, R., Magnusson, W.E., Martin, C.S., McGeoch, M., Mihoub, J.-B., Pettorelli, N., Proença, V., Peng, C., Regan, E., Schmiedel, U., Simaika, J.P., Weatherdon, L., Waterman, C., Xu, H., Belnap, J., 2017. Building capacity in biodiversity monitoring at the global scale. Biodivers Conserv 26, 2765–2790. https://doi.org/10.1007/s10531-017-1388-7

Schowalter, T., 2017. Arthropod Diversity and Functional Importance in Old-Growth Forests of North America. Forests 8, 97. https://doi.org/10.3390/f8040097

Sekercioglu, C.H., 2012. Bird functional diversity and ecosystem services in tropical forests, agroforests and agricultural areas. J Ornithol 153, 153–161. https://doi.org/10.1007/s10336-012-0869-4

Şekercioğlu, Ç.H., Primack, R.B., Wormworth, J., 2012. The effects of climate change on tropical birds. Biol Conserv 148, 1–18. https://doi.org/10.1016/j.biocon.2011.10.019

Sheeren, D., Bonthoux, S., Balent, G., 2014. Modeling bird communities using unclassified remote sensing imagery: Effects of the spatial resolution and data period. Ecol Indic 43, 69–82. https://doi.org/10.1016/j.ecolind.2014.02.023

Silva Junior, C.H.L., Pessôa, A.C.M., Carvalho, N.S., Reis, J.B.C., Anderson, L.O., Aragão, L.E.O.C., 2020. The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. Nat Ecol Evol 5, 144–145. https://doi.org/10.1038/s41559-020-01368-x

Simons, T.R., Alldredge, M.W., Pollock, K.H., Wettroth, J.M., 2007. Experimental Analysis of The Auditory Detection Process on Avian Point Counts. Auk 124, 986–999. https://doi.org/10.1093/auk/124.3.986

Singh, A., Meena, G.K., Kumar, S., Gaurav, K., 2019. Evaluation of the Penetration Depth of L-and S-Band (NISAR mission) Microwave SAR Signals into Ground.

Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali, A., Schaepman, M.E., Paganini, M., Kissling, W.D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., Fernandez, M., Fernández, N., Gorelick, N., Geijzendorffer, I., Heiden, U., Heurich, M., Hobern, D., Holzwarth, S., Muller-Karger, F.E., Van De Kerchove, R., Lausch, A., Leitão, P.J., Lock, M.C., Mücher, C.A., O’Connor, B., Rocchini, D., Roeoesli, C., Turner, W., Vis, J.K., Wang, T., Wegmann, M., Wingate, V., 2021. Priority list of biodiversity metrics to observe from space. Nat Ecol Evol 5, 896–906. https://doi.org/10.1038/s41559-021-01451-x

Skidmore, A.K., Pettorelli, N., Coops, N.C., Geller, G.N., Hansen, M., Lucas, R., Mücher, C.A., O’Connor, B., Paganini, M., Pereira, H.M., Schaepman, M.E., Turner, W., Wang, T., Wegmann, M., 2015. Environmental science: Agree on biodiversity metrics to track from space. Nature 523, 403–405. https://doi.org/10.1038/523403a

Slingsby, J.A., Merow, C., Aiello-Lammens, M., Allsopp, N., Hall, S., Mollmann, H.K., Turner, R., Wilson, A.M., Silander, J.A., 2017. Intensifying postfire weather and biological invasion drive species loss in a Mediterranean-type biodiversity hotspot. Proc Natl Acad Sci U S A 114, 4697–4702. https://doi.org/10.1073/pnas.1619014114

Smith, C.C., Espírito-Santo, F.D.B., Healey, J.R., Young, P.J., Lennox, G.D., Ferreira, J., Barlow, J., 2020. Secondary forests offset less than 10% of deforestation-mediated carbon emissions in the Brazilian Amazon. Glob Chang Biol 26, 7006–7020. https://doi.org/10.1111/gcb.15352

Smith, P., Bustamante, M., Ahammad, H., Clark, H., Dong, H., Elsiddig, E.A., 2014. Agriculture, Forestry and Other Land Use (AFOLU). Climate Change 2014: Mitigation of Climate Change 811–922.

Solberg, S., Næsset, E., Gobakken, T., Bollandsås, O.-M., 2014. Forest biomass change estimated from height change in interferometric SAR height models. Carbon Balance Manag 9, 5. https://doi.org/10.1186/s13021-014-0005-2

Steenweg, R., Hebblewhite, M., Whittington, J., McKelvey, K., 2019. Species-specific differences in detection and occupancy probabilities help drive ability to detect trends in occupancy. Ecosphere 10. https://doi.org/10.1002/ecs2.2639

Stein, A., Gerstner, K., Kreft, H., 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecol Lett 17, 866–880. https://doi.org/10.1111/ele.12277

Stephenson, P., 2020. Technological advances in biodiversity monitoring: applicability, opportunities and challenges. Curr Opin Environ Sustain 45, 36–41. https://doi.org/10.1016/j.cosust.2020.08.005

Stork, N.E., Boyle, T.J.B., Dale, V., Eeley, H., Finegan, B., Lawes, M., Manokaran, N., Prabhu, R., J., S., 1997. Criteria and indicators for assessing the sustainability of forest management: conservation of biodiversity, Criteria and indicators for assessing the sustainability of forest management: conservation of biodiversity. Center for International Forestry Research (CIFOR). https://doi.org/10.17528/cifor/000077

Stork, N.E., McBroom, J., Gely, C., Hamilton, A.J., 2015. New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. Proceedings of the National Academy of Sciences 112, 7519–7523. https://doi.org/10.1073/pnas.1502408112

Stouffer, P.C., Johnson, E.I., Bierregaard, R.O., Lovejoy, T.E., 2011. Understory Bird Communities in Amazonian Rainforest Fragments: Species Turnover through 25 Years Post-Isolation in Recovering Landscapes. PLoS One 6, e20543. https://doi.org/10.1371/journal.pone.0020543

Stratford, J.A., Stouffer, P.C., 2015. Forest fragmentation alters microhabitat availability for Neotropical terrestrial insectivorous birds. Biol Conserv 188, 109–115. https://doi.org/10.1016/j.biocon.2015.01.017

Subedi, B., Poudel, A., Aryal, S., 2023. The impact of climate change on insect pest biology and ecology: Implications for pest management strategies, crop production, and food security. J Agric Food Res 14, 100733. https://doi.org/10.1016/j.jafr.2023.100733

Suttidate, N., Hobi, M.L., Pidgeon, A.M., Round, P.D., Coops, N.C., Helmers, D.P., Keuler, N.S., Dubinin, M., Bateman, B.L., Radeloff, V.C., 2019. Tropical bird species richness is strongly associated with patterns of primary productivity captured by the Dynamic Habitat Indices. Remote Sens Environ 232, 111306. https://doi.org/10.1016/j.rse.2019.111306

Terraube, J., Archaux, F., Deconchat, M., van Halder, I., Jactel, H., Barbaro, L., 2016. Forest edges have high conservation value for bird communities in mosaic landscapes. Ecol Evol 6, 5178–5189. https://doi.org/10.1002/ece3.2273

Thiel, C., Drezet, P., Weise, C., Quegan, S., Schmullius, C., 2006. Radar remote sensing for the delineation of forest cover maps and the detection of deforestation. Forestry 79, 589–597. https://doi.org/10.1093/forestry/cpl036

Thornton, D.H., Branch, L.C., Sunquist, M.E., 2011. The influence of landscape, patch, and within-patch factors on species presence and abundance: a review of focal patch studies. Landsc Ecol 26, 7–18. https://doi.org/10.1007/s10980-010-9549-z

Tikhonov, G., Ovaskainen, O., Oksanen, J., de Jonge, M., Opedal, O., Dallas, T., 2021. Hmsc: Hierarchical Model of Species Communities.

Titeux, N., Henle, K., Mihoub, J., Regos, A., Geijzendorffer, I.R., Cramer, W., Verburg, P.H., Brotons, L., 2016. Biodiversity scenarios neglect future land-use changes. Glob Chang Biol 22, 2505–2515. https://doi.org/10.1111/gcb.13272

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens Environ 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

Torresani, M., Rocchini, D., Sonnenschein, R., Zebisch, M., Marcantonio, M., Ricotta, C., Tonon, G., 2019. Estimating tree species diversity from space in an alpine conifer forest: The Rao's Q diversity index meets the spectral variation hypothesis. Ecol Inform 52, 26–34. https://doi.org/10.1016/j.ecoinf.2019.04.001

Traylor, C.R., Ulyshen, M.D., Wallace, D., Loudermilk, E.L., Ross, C.W., Hawley, C., Atchison, R.A., Williams, J.L., McHugh, J. V., 2022. Compositional attributes of invaded forests drive the diversity of insect functional groups. Glob Ecol Conserv 35, e02092. https://doi.org/10.1016/J.GECCO.2022.E02092

Tsai, F., Lin, E. -K., Yoshino, K., 2007. Spectrally segmented principal component analysis of hyperspectral imagery for mapping invasive plant species. Int J Remote Sens 28, 1023–1039. https://doi.org/10.1080/01431160600887706

Tu, H.-M., Fan, M.-W., Ko, J.C.-J., 2020. Different Habitat Types Affect Bird Richness and Evenness. Sci Rep 10, 1221. https://doi.org/10.1038/s41598-020-58202-4

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., Woodcock, C., 2015. Free and open-access satellite data are key to biodiversity conservation. Biol Conserv 182, 173–176. https://doi.org/10.1016/j.biocon.2014.11.048

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends Ecol Evol 18, 306–314. https://doi.org/10.1016/S0169-5347(03)00070-3

Tyukavina, A., Potapov, P., Hansen, M.C., Pickens, A.H., Stehman, S. V., Turubanova, S., Parker, D., Zalles, V., Lima, A., Kommareddy, I., Song, X.-P., Wang, L., Harris, N., 2022. Global Trends of Forest Loss Due to Fire From 2001 to 2019. Frontiers in Remote Sensing 3. https://doi.org/10.3389/frsen.2022.825190

Uhler, J., Haase, P., Hoffmann, L., Hothorn, T., Schmidl, J., Stoll, S., Welti, E.A.R., Buse, J., Müller, J., 2022. A comparison of different Malaise trap types. Insect Conserv Divers 15, 666–672. https://doi.org/10.1111/icad.12604

UNESCO, 2021. Cape Floral Region Protected Areas - UNESCO World Heritage Centre [WWW Document]. URL https://whc.unesco.org/en/list/1007/ (accessed 7.19.22).

UNFCCC, 2023. Home / REDD+ [WWW Document]. United Nations Framework Convention on Climate Change. URL https://redd.unfccc.int/ (accessed 11.8.23).

United Nations, 2021. A New Global Framework for Managing Nature Through 2030: 1st Detailed Draft Agreement Debuts – United Nations Sustainable Development [WWW Document]. URL https://www.cbd.int/article/draft-1-global-biodiversity-framework (accessed 5.13.22).

Ustin, S.L., Jacquemoud, S., 2020. How the optical properties of leaves modify the absorption and scattering of energy and enhance leaf functionality, in: Remote Sensing of Plant Biodiversity. Springer International Publishing, pp. 349–384. https://doi.org/10.1007/978-3-030-33157-3_14

van der Maarel, E., 2007. Transformation of cover-abundance values for appropriate numerical treatment - Alternatives to the proposals by Podani. Journal of Vegetation Science 18, 767–770. https://doi.org/10.1111/j.1654-1103.2007.tb02592.x

van der Merwe, H., 2020. Vegetation of the Square Kilometre Array (SKA), South Africa: A baseline to measure local and global change. South African Journal of Botany 132, 22–29. https://doi.org/10.1016/j.sajb.2020.03.025

van der Merwe, H., Hoffman, T.M., 2019. Vegetation of Akkerendam Nature Reserve, Northern Cape: Delineation and dynamics over 100 years. Bothalia. https://doi.org/10.4102/abc.v49i1.2401

Van der Merwe, H., Van Rooyen, M.W., Van Rooyen, N., 2008a. Vegetation of the Hantam-Tanqua-Roggeveld subregion, South Africa. Part 1: Fynbos Biome related vegetation. Koedoe 50, 61–71. https://doi.org/10.4102/koedoe.v50i1.130

Van der Merwe, H., Van Rooyen, M.W., Van Rooyen, N., 2008b. Vegetation of the Hantam-Tanqua-Roggeveld subregion, South Africa Part 2: Succulent Karoo Biome related vegetation. Koedoe 50, 160–183. https://doi.org/10.4102/koedoe.v50i1.148

van Emmerik, T., Steele-Dunne, S., Paget, A., Oliveira, R.S., Bittencourt, P.R.L., Barros, F. de V., van de Giesen, N., 2017. Water stress detection in the Amazon using radar. Geophys Res Lett 44, 6841–6849. https://doi.org/10.1002/2017GL073747

van Wees, D., van der Werf, G.R., Randerson, J.T., Andela, N., Chen, Y., Morton, D.C., 2021. The role of fire in global forest loss dynamics. Glob Chang Biol 27, 2377–2391. https://doi.org/10.1111/gcb.15591

Vergara, A., Arias, M., Gachet, B., Naranjo, L.G., Román, L., Surkin, J., Tamayo, V., 2022. Living Amazon Report 2022.

Vittinghoff, E., McCulloch, C.E., 2007. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. Am J Epidemiol 165, 710–718. https://doi.org/10.1093/aje/kwk052

Vlok, J., Cowling, R., Wolf, T., 2005. A Vegetation Map for the Little Karoo.

Wang, G., Cai, W., Gan, B., Wu, L., Santoso, A., Lin, X., Chen, Z., McPhaden, M.J., 2017. Continued increase of extreme El Niño frequency long after 1.5 °C warming stabilization. Nat Clim Chang 7, 568–572. https://doi.org/10.1038/nclimate3351

Wang, R., Gamon, J.A., 2019. Remote sensing of terrestrial plant biodiversity. Remote Sens Environ 231, 111218. https://doi.org/10.1016/j.rse.2019.111218

Wang, R., Gamon, J.A., Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Zygielbaum, A.I., Kothari, S., 2018. Influence of species richness, evenness, and composition on optical diversity: A simulation study. Remote Sens Environ 211, 218–228. https://doi.org/10.1016/j.rse.2018.04.010

Warren, Dan.L., Cardillo, M., Rosauer, D.F., Bolnick, D.I., 2014. Mistaking geography for biology: inferring processes from species distributions. Trends Ecol Evol 29, 572–580. https://doi.org/10.1016/j.tree.2014.08.003

Warren, S.D., Alt, M., Olson, K.D., Irl, S.D.H., Steinbauer, M.J., Jentsch, A., 2014. The relationship between the spectral diversity of satellite imagery, habitat heterogeneity, and plant species richness. Ecol Inform 24, 160–168. https://doi.org/10.1016/J.ECOINF.2014.08.006

Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. Journal of Machine Learning Research 11, 3571–3594.

West, T.A.P., Börner, J., Sills, E.O., Kontoleon, A., 2020. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. Proceedings of the National Academy of Sciences 117, 24188–24194. https://doi.org/10.1073/pnas.2004334117

Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R., McCarthy, M.A., 2021. Defining and evaluating predictions of joint species distribution models. Methods Ecol Evol 12, 394–404. https://doi.org/10.1111/2041-210X.13518

Williams, E.M., O'Donnell, C.F.J., Armstrong, D.P., 2018. Cost-benefit analysis of acoustic recorders as a solution to sampling challenges experienced monitoring cryptic species. Ecol Evol 8, 6839–6848. https://doi.org/10.1002/ece3.4199

Willis, K.J., Araújo, M.B., Bennett, K.D., Figueroa-Rangel, B., Froyd, C.A., Myers, N., 2007. How can a knowledge of the past help to conserve the future? Biodiversity conservation and the relevance of long-term ecological studies. Philosophical

Transactions of the Royal Society B: Biological Sciences 362, 175–187. https://doi.org/10.1098/rstb.2006.1977

Willmer, J.N.G., Püttker, T., Prevedello, J.A., 2022. Global impacts of edge effects on species richness. Biol Conserv 272, 109654. https://doi.org/10.1016/j.biocon.2022.109654

Wimmer, J., Towsey, M., Roe, P., Williamson, I., 2013. Sampling environmental acoustic recordings to determine bird species richness. Ecological Applications 23, 1419–1428. https://doi.org/10.1890/12-2088.1

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., 2008. Effects of sample size on the performance of species distribution models. Divers Distrib 14, 763–773. https://doi.org/10.1111/j.1472-4642.2008.00482.x

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.-A., Guisan, A., Heikkinen, R.K., Høye, T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.A., Aastrup, P., Svenning, J.-C., 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. OnlineOpen_Terms Biological Reviews 88, 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x

Wolda, H., 1987. Altitude, habitat and tropical insect diversity. Biological Journal of the Linnean Society 30, 313–323. https://doi.org/10.1111/j.1095-8312.1987.tb00305.x

Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., Zhu, Z., 2019. Current status of Landsat program, science, and applications. Remote Sens Environ 225, 127–147. https://doi.org/10.1016/J.RSE.2019.02.015

Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. Remote Sens Environ 185, 271–283. https://doi.org/10.1016/j.rse.2015.11.032

Wunder, S., Schulz, D., Montoya-Zumaeta, J.G., Börner, J., Ponzoni Frey, G., Betancur-Corredor, B., 2024. Modest forest and welfare gains from initiatives for reduced emissions from deforestation and forest degradation. Commun Earth Environ 5, 394. https://doi.org/10.1038/s43247-024-01541-1

WWF-Brazil, 2024. In one year, deforestation and conversion falls 30.6% in the Amazon and 25.7% in the Cerrado | WWF Brasil [WWW Document]. URL https://www.wwf.org.br/nossosconteudos/notas_e_releases/english/?90200/In-one-year-deforestation-falls-306-in-the-Amazon-and-257-in-the-Cerrado (accessed 2.14.25).

Xing, F., An, R., Wang, B., Miao, J., Jiang, T., Huang, X., Hu, Y., 2021. Mapping the occurrence and spatial distribution of noxious weed species with multisource data in degraded grasslands in the Three-River Headwaters Region, China. Science of The Total Environment 801, 149714. https://doi.org/10.1016/j.scitotenv.2021.149714

Xu, R., Li, Y., Teuling, A.J., Zhao, L., Spracklen, D. V., Garcia-Carreras, L., Meier, R., Chen, L., Zheng, Y., Lin, H., Fu, B., 2022. Contrasting impacts of forests on cloud cover based on satellite observations. Nat Commun 13, 670. https://doi.org/10.1038/s41467-022-28161-7

Xu, W.B., Svenning, J.C., Chen, G.K., Zhang, M.G., Huang, J.H., Chen, B., Ordonez, A., Ma, K.P., 2019. Human activities have opposing effects on distributions of narrow-ranged and widespread plant species in China. Proc Natl Acad Sci U S A 116, 26674–26681. https://doi.org/10.1073/pnas.1911851116

Xue, J., Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. J Sens 2017, 1–17. https://doi.org/10.1155/2017/1353691

Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C., Ding, Z., 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods Ecol Evol 3, 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zhang, J., Nielsen, S.E., Grainger, T.N., Kohler, M., Chipchar, T., Farr, D.R., 2014. Sampling Plant Diversity and Rarity at Landscape Scales: Importance of Sampling Time in Species Detectability. PLoS One 9, e95334. https://doi.org/10.1371/journal.pone.0095334

Zhang, K., Lin, S., Ji, Y., Yang, Chenxue, Wang, X., Yang, Chunyan, Wang, H., Jiang, H., Harrison, R.D., Yu, D.W., 2016. Plant diversity accurately predicts insect diversity in two tropical landscapes. Mol Ecol. https://doi.org/10.1111/mec.13770

Zhang, L., Ji, L., Wylie, B.K., 2011. Response of spectral vegetation indices to soil moisture in grasslands and shrublands. Int J Remote Sens 32, 5267–5286. https://doi.org/10.1080/01431161.2010.496471

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens Environ 84, 471–475. https://doi.org/10.1016/S0034-4257(02)00135-9

Zhong, B., Yang, L., Luo, X., Wu, J., Hu, L., 2024. Extracting Shrubland in Deserts from Medium-Resolution Remote-Sensing Data at Large Scale. Remote Sens (Basel) 16, 374. https://doi.org/10.3390/rs16020374

Zhou, J., Yan Guo, R., Sun, M., Di, T.T., Wang, S., Zhai, J., Zhao, Z., 2017. The Effects of GLCM parameters on LAI estimation using texture values from Quickbird Satellite Imagery. Sci Rep 7, 7366. https://doi.org/10.1038/s41598-017-07951-w

Zvoleff, A., 2020. glcm: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs).

7. Appendices

7.1 Appendix - Chapter 3

DNA was extracted from each bulk tissue sample using a proprietary lysis buffer and a DNeasy Blood and Tissue Kit (Qiagen). First, ethanol was poured off and samples were left to dry for 24 hours in an incubator at 30°C. Enough lysis buffer was added to cover each sample (200-400 ml), then samples were incubated at 56°C for 4 hours. After incubation, 14 ml of lysate from each sample was passed through a set of five spin columns (2.8 ml per spin column), followed by 500 µl of Buffer AW1, then 500 µl of Buffer AW2. Finally, 50 µl of Buffer AE was passed through each spin column and the five extracts per sample were pooled together.

An extraction blank was processed with each batch of extractions to assess potential contamination in the extraction process. DNA extracts were quantified using a Qubit dsDNA HS Assay Kit on a Qubit 3.0 fluorometer (Thermo Scientific).

For each bulk invertebrate sample, three replicates were performed using COI primers (Wangensteen et al., 2018). All PCRs were carried out in a total volume of 25 µL consisting of 1X DreamTaq Green PCR Master Mix, 0.3 µM of each primer, 0.8 mg/ml of BSA, 0.9 µl of template DNA, and PCR grade water.

Bulk invertebrate PCR conditions consisted of: an initial denaturation at 95°C for 10 min; 35 cycles at 94°C for 60 s, 45°C for 60 s, and 72°C for 60 s; and a final elongation step at 72°C for 5 min. PCR positive controls (i.e. a mock community with a known composition of non-native species) were included to verify sequence quality and PCR negative controls (i.e. PCR grade water) were included to detect potential cross-contamination. Amplification success was confirmed via gel electrophoresis.

Library preparation. All first round PCR replicates per sample per marker were pooled and purified using Mag-Bind® TotalPure NGS (Omega Bio-tek) magnetic beads. A sequencing library was prepared from the purified amplicons using a combinational dual index approach, following Illumina's 16S Metagenomic Sequencing Library Preparation protocol using the Nextera XT indexes (Illumina), but using 1X DreamTaqPCR Master Mix (Thermo Scientific). The second round indexed PCR products were again purified using Mag-Bind® TotalPure NGS (Omega Bio-tek) magnetic beads. The purified index products were quantified using a Qubit dsDNA BR Assay Kit, normalized to 4 nM and pooled at equal volumes. The pooled 4 nM purified index PCRs were sized using a TapeStation D1000 ScreenTape System (Agilent). The libraries were sequenced on an Illumina MiSeq with a V3 MiSeq Reagent kit, the final library was loaded at 10 pM with a 20% PhiX control spike.

Sequences were demultiplexed with bcl2fastq and subsequently processed to obtain Amplicon Sequence Variants (ASV).

Taxonomic assignments were made via sequence similarity (blastn; (Altschul et al., 1990; Camacho et al., 2009) searches of the ASV sequences against the NCBI nucleotide database (NCBI nt; downloaded on 28-09-2021) and the BOLD datasets (Ratnasingham and Hebert, 2007). Hits were required to have a minimum e-score of 1e-20 and cover at least 90% of the query sequence. The taxonomy associated with each hit was converted to the GBIF taxonomic backbone to allow results from different databases to be comparable for the following step.

Public databases are known to contain errors with DNA sequences frequently associated with incorrect species names. This poses a problem when assigning species names to metabarcoding DNA sequences. NatureMetrics uses proprietary methods to curate and filter out such errors through whitelisting, blacklisting, and human curation based on country lookups to ensure that there is reasonable support that the sequence has been correctly associated in the database. Accessions that are flagged as potentially erroneous are not used for taxonomic assignment of metabarcoding sequences. It should be noted that while it is possible to use uncurated public databases for taxonomic assignment, resulting outputs might differ (slightly), such as in species resolution, from those resulting from curated ones.

ASVs were then clustered at 97% similarity with USEARCH to obtain OTUs. An OTU-by-sample table was generated by mapping all dereplicated reads for each sample to the OTU representative sequences with USEARCH at an identity threshold of 97%.

Finally, low abundance detections were omitted, with filter thresholds set at a percentage of the total reads per sample (vertebrate eDNA: 0.02%; invertebrate eDNA: 0.015%; bacteria: 0.1%; bulk invertebrates: 0.01%). Results are presented for OTUs identified to the target taxonomic group only (vertebrate: Chordata excluding human and domestic animals; invertebrate: Animalia excluding Tetrapoda; bacterial: Bacteria).

Sequencing data are available via this link.

https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdataview.ncbi.nlm.nih.gov%2Fobject%2FPRJNA898889%3Freviewer%3Dl9cfd6aa3g5aluolad91qtcit5&data=05%7C01%7Cslatera6%40live.lancs.ac.uk%7C3ff7c5cb0c0d43cd572308dba49a1b2a%7C9c9bcd11977a4e9ca9a0bc734090164a%7C0%7C0%7C638284754010652951%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=iI5BwaBRzup5NEj3kbdzhAKFlrG9xO3bOniUl6S2IBQ%3D&reserved=0