BIODIVERSITY MONITORING, EARTH OBSERVATIONS AND THE ECOLOGY

OF SCALE

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF BIOLOGY

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Christopher Benjamin Anderson

December 2020

This dissertation is online at: http://purl.stanford.edu/pd244cm6715

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Gretchen Daily, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Chris Field

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Erin Mordecai

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

Climate change, agricultural expansion, and population growth are dramatically altering ecosystem structure, function, and community composition worldwide. Yet our ability to measure, monitor, and forecast biodiversity change—crucial for mitigating it—remains limited. Until recently, global biodiversity monitoring systems seemed far-fetched. The sheer amount of data that would be required to map the variation in genes, species, communities, and ecosystems that comprises biodiversity is enormous, to say nothing of mapping change over time. Until recently, ecology was perceived as a data-scarce discipline, with field plots and species occurrence records sparsely and opportunistically collected, and access fragmented among research groups. But several advances in the past decade—the deployment of hundreds of earth observing sensors, the consolidation and standardization of open ecological data, and increased access to open software and computing resources—have significantly reduced the ecological data gap that biodiversity monitoring systems needed to bridge. What should we make of all this new information? Right now, the data alone are insufficient for monitoring purposes. Ecological data are subject to taxonomic, geographical, or recency biases, and satellites do not record measurements in ecological units; they're often in units of energy (e.g., $W\,m^{-2}\,sr^{-1}\,\mu m^{-1}$). But a framework that links these disparate data types—discrete, spatially-explicit species data and continuous, feature-rich, and regularly-updated earth observations imagery—could be used to precisely map species locations and habitat across large extents over time. In this dissertation, I demonstrate how concepts of pattern and scale in ecology could be used to design such a framework, as these concepts apply to both ecological and to earth observations analyses.

Following a literature review of the scale-dependent challenges to linking *in situ* and earth observations data, I describe two scale-explicit machine learning approaches to species mapping: classifying species identities for individual tree crowns in high resolution airborne earth observations

data, and mapping the niches and distributions of two mosquito arbovirus vectors, *Aedes aegypti* and *Ae. albopictus*, using low resolution satellite data. The tree species modeling approach identified crowns with $> 90\%$ accuracy, enabling precision species mapping over large areas. The mosquito modeling approach identified novel drivers of the spatial distributions for these arbovirus vectors, revealing that resource constraints (i.e., access to blood meals) play a central role in determining distribution patterns for these globally invasive species. These applications span several orders of biological magnitude, mapping some of the largest (trees) and smallest (mosquitoes) terrestrial macro-organisms across landscape and continental extents. This work advances the conceptual and technical basis for deploying satellite-based biodiversity monitoring systems using ecological scaling principles.

Acknowledgments

This work was deeply influenced by conversations and friendships with many kind and smart people, to whom I owe so much.

Gretchen Daily welcomed me to her lab in the Biology Department in early 2016. She has been a champion and an advocate for me every day since, and I am profoundly grateful for the opportunities she has created for me. I am inspired by her commitment to understanding the value of nature—from all perspectives—and to finding opportunities to conserve and cherish our planet's wondrous biodiversity. We've always shared the commitment; she taught me how to translate it into action. I've learned so much from Gretchen: how to communicate my work with passion and precision; how to build an organization dedicated to inclusive, science-driven conservation; the value of leading through empowering others; and how to meet every person where they are. As my primary adviser, Gretchen was solely focused on my development as a scientist, ensuring I was working on problems I was passionate about and that I had access to the people and resources that would help me grow. I benefit tremendously as a result. Thank you.

I have had the honor and the privilege of being advised by Erin Mordecai and Chris Field, two thoughtful, critical and generous professors who inspire and challenge me in every discussion we have. When most of my early research focused on finding empirical patterns in ecological data, Erin encouraged me to dig deeper and deeper into the biological mechanisms that might underpin them, helping me think more clearly and test assumptions more rigorously than I would have otherwise. Chris has a wealth of knowledge and experience as a scientist and as an adviser. I am regularly amazed by his ability to ask the right question or provide the right recommendation to help me think through an issue more clearly and guiding me towards novel research frontiers. They have both shaped my perspective as a scientist, and I am grateful for their advice.

surveying forests from the sky. I've flown transects from the Andes to the Amazon, witnessing both the unparalleled beauty of the transition from treeline to the lowlands, and the rapid loss of forests to fires, gold mining, and logging. The experience richly colored my understanding of the earth system: as beautiful, complex, very big, and changing in ways that are hard to convey with numbers alone.

I was also fortunate to meet and learn from an outstanding group of scientists during my time at the Department of Global Ecology. This includes Claire Baldeck, Jomar Barbosa, Joe Berry, Joe Boardman, Paulo Brando, Phil Brodrick, Loreli Carranza-Jiménez, Dana Chadwick, Cecilia Chavana-Bryant, John Clark, Andrew Davies, Shane Easter, Michael Eastwood, Jean-Baptiste Féret, Emily Francis, Mark Helmlinger, Mona Houcheime, James Jacobson, Jen Johnson, Ty Kennedy-Bowdoin, Dave Knapp, Dave Marvin, Joe Mascaro, Kelly McManus, Elsa Ordway, Ted Raab, Elif Tassar, Phil Taylor, Todd Tobeck, Raul Tupayachi, Nick Vaughn, and Parker Weiss. I owe yet more thanks to Chris Field, then the director of the department, who fostered an open and welcoming environment where I was empowered to learn and grow.

Thank you to my parents, Rhonda and Robert Anderson, for loving me and supporting my passion for ecology. You've taught me so much, and I cherish you both.

I am grateful for my wife Helen, who shared her kindness and her love with me before and throughout my time at Stanford. She taught me a very important lesson soon after we met: that interested people are interesting. I reflect on this often, which reminds me that listening is both an expression of kindness and caring for others as well as a means to grow as a person. She reinforces my appreciation of libraries through her work, and demonstrates an inspiring example of the value of service to others. Thank you for sharing so much with me.

Epigraph

Bertrand Russell declared that, in case he met God, he would say to [them], "You did not give us enough information." I would add to that, "All the same, I'm not persuaded that we did the best we could with the information we had. Toward the end there, anyway, we had tons of information."

Kurt Vonnegut Jr., Palm Sunday

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Christopher B. Anderson

1.1 Global Change Ecology

The form and function of the earth's ecosystems are rapidly changing. Agricultural lands are now the largest terrestrial biome in the world, occupying approximately 40% of the land surface [Foley et al., 2005, Springmann et al., 2018]. This expansion occurred at the expense of the world's forests, exerting an exacting toll in the tropics. There was a net loss of 42 million hectares of forest between 2000 and 2010 in tropical countries, commensurate with a net gain of 36 million hectares in agricultural lands [FAO, 2016]. Roughly two-thirds of agricultural area is pasture for livestock, which comprises approximately 60% of global mammalian biomass, significantly outweighing humans and wildlife, which comprise the remaining 36% and 4% [Ramankutty et al., 2008, Bar-On et al., 2018]. Simultaneously, forested area is increasing across temperate and boreal systems, and it is unclear whether total global forest area is increasing or decreasing over decadal time scales [Hansen et al., 2013, Song et al., 2018].

This restructuring of the earth's ecosystems could be framed as a global game of whack-a-mole. Wildlife populations are crashing while livestock populations are booming. Forests are cut and cleared in the tropics, replanted and restored in temperate and boreal systems (Fig. 1.1). This has had tremendous effects on terrestrial biodiversity; around 25% of plant and animal groups are

Figure 1.1: Transition from forest to agriculture at the edge of the Las Alturas Wildlife Sanctuary in Costa Rica. Land use change is fundamentally altering patterns of ecosystem structure and ecosystem function, including the amount of wildlife habitat and patterns of primary productivity.

threatened, the extinction rate is now tens to hundreds of times higher than the average rate over the past 10 million years, and the number of invasive species per-country has increased nearly 70%, all leading to the widespread erosion of differences between ecological communities [IPBES, 2019]. Ecosystems are a canonical example of complex, interconnected systems, and these changes are bound to have far-reaching consequences for other ecological processes. This includes the transmission of vector-borne diseases, which are poised to increase in frequency and intensity with environmental change, potentially exposing an additional billion people to the threat of Zika alone [Ryan et al., 2020]. With so much at stake, biodiversity monitoring systems are now being developed to track these complex patterns—the winners, the losers, the just differents—and to mitigate the effects of change.

A series of biodiversity monitoring frameworks have recently been developed to systematically

assess change for multiple taxa over large extents [Scholes et al., 2012, Fernández et al., 2015]. These monitoring frameworks have been shaped in large part by open access to biodiversity data [Kattge et al., 2011, Jetz et al., 2012, Metzger et al., 2013, Culina et al., 2018] as well as access to an array of modeling tools to analyze these data [Butchart et al., 2010, Pettorelli et al., 2016, Gorelick et al., 2017]. Most contemporary monitoring frameworks map biodiversity patterns according to the Essential Biodiversity Variables framework, a hierarchical grouping of metrics that quantify the variation in genes, species, communities and ecosystems [Pereira et al., 2013]. Inspired by the Essential Climate Variables framework, these metrics are i) biological, ii) sensitive to change, and iii) ecosystem agnostic, meaning they can be mapped anywhere. This definition is broad enough to include patterns of ecosystem structure, like aboveground biomass or habitat fragmentation, which are not often referred to as biodiversity *per se*. This flexible and inclusive framework—which extends beyond a narrower definition of biodiversity as patterns of species richness, evenness and abundance—is now the standard for global monitoring systems, as it has been adopted by the intergovernmental Group on Earth Observations [GEO BON, 2017].

While the amount of open biodiversity data is now nothing short of incredible—the Global Biodiversity Information Facility (`https://www.gbif.org`) now hosts over 1.6 billion unique species occurrence records—the data gaps faced by monitoring systems are too wide to be bridged by *in situ* biodiversity data alone. The spatial, taxonomic and temporal gaps in open biodiversity data are well documented [Beck et al., 2014, Hortal et al., 2015, Geijzendorffer et al., 2016], and a central task for the global biodiversity mapping community is to develop modeling approaches that address bias issues and extrapolate from incomplete records to map large extents.

Earth observations data, especially satellite imagery, are a natural compliment to *in situ* data because they provide consistent measurements of terrestrial ecosystems over time, characterising biodiversity patterns in accessible and inaccessible areas alike. And plenty of data are available: at least 44 satellite constellations have been launched by 20 space agencies since 1977, which have been used to measure at least 15 different biodiversity patterns (Table B.1). But using satellites to map biodiversity is not a straightforward task. Satellites measure patterns of electromagnetic radiation, like spectral radiance, thermal emmissivity, or microwave backscattering, which are not in biological units of measurement. Given this fundamental difference in data types, it is worth reviewing the technical and conceptual basis for mapping biodiversity from earth observing sensors. How are

satellite measurements converted from units of energy into metrics that are biological, sensitive to change, and ecosystem agnostic?

1.2 Biodiversity Mapping with Earth Observations Data

The majority of earth observing satellites are passive optical sensors that measure the amount of solar irradiance reflected by the earth's surface, typically in the spectrum of visible to shortwave infrared light (0.4-2.5 μm). Terrestrial vegetation, which covers the majority of the land surface, interacts with light in three ways. It absorbs light, with photosynthetic pigments harvesting photons to drive the electron transport chain and generate ATP. Vegetation also transmits light, often as a photoprotective mechanism to reduce potential molecular damage from harmful infrared radiation. The remaining light is reflected. This can be bidirectional, reflecting energy at an equal and opposite angle to the light source and normal to a leaf's surface, or it could be refracted and scattered spherically through the leaf tissue. The amount of light absorbed, transmitted, or reflected varies by wavelength and varies in response to many biophysical factors. Since most earth observing sensors measure reflected light, it is critical to understand the ecological processes driving vegetation reflectance patterns. Vegetation reflectance is driven by four general factors:

1. Leaf biochemistry, including the concentrations of pigments and defense compounds (Fig. 1.2A).

2. Plant resource use, including leaf water content and nitrogen content (Fig. 1.2B).

3. Leaf structure, including cell wall composition (e.g., cellulose and lignin concentrations) and the amount of intracellular air space (Fig. 1.2C).

4. Canopy structure, including the number and orientation of leaves on a plant (Fig. 1.2D).

This is also important in the context of biodiversity monitoring because these four factors are also indicators for discriminating between species. Plant biochemistry is well-conserved among individuals of the same species [Kokaly et al., 2009, Asner et al., 2011, Funk et al., 2017], as are resource use patterns and leaf and canopy structure patterns [Hallé et al., 1978, Townsend et al., 2007]. Interspecific variation in these patterns is driven by resource access, competition and co-evolution, leading to niche partitioning and differentiation along multiple niche axes [Wright et al., 2004, Osnas et al., 2013]. Reflectance patterns therefore contain information that could be used to classify
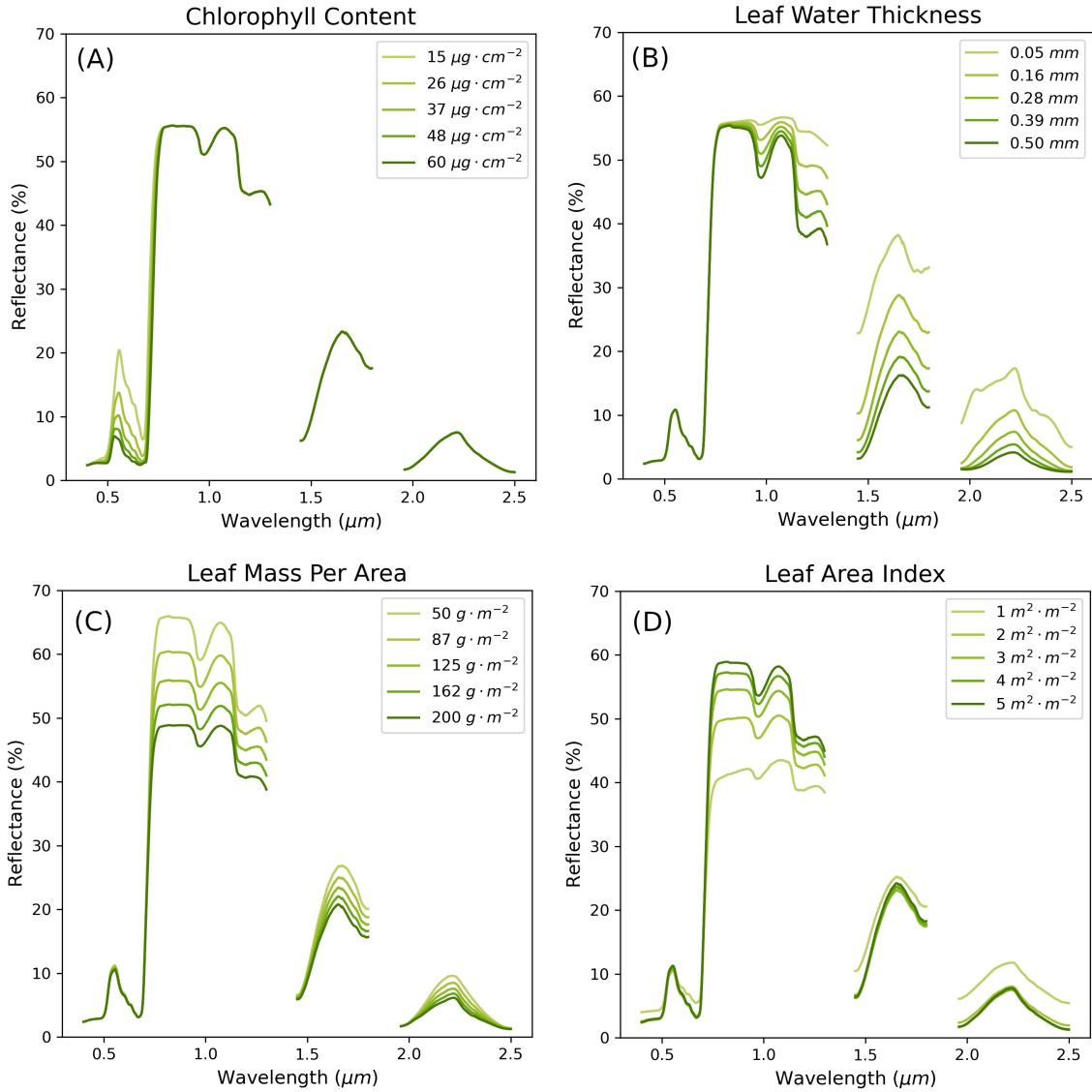
Figure 1.2: Simulated vegetation reflectance spectra showing the nonlinear effects of variation in four leaf and canopy properties: (A) leaf chlorophyll content, (B) leaf equivalent water thickness, (C), leaf mass per area, and (D) leaf area index. All spectra were generated using the PROSAIL coupled leaf and canopy model [Jacquemoud et al., 2009] using a range of global parameter values [Rivera et al., 2013]. For (A)-(D), each parameter was set to the global mean, with the exception of the focal parameter, which varied in a range from min and max of the 95% range. Bands dominated by atmospheric water vapor concentrations (1.3-1.45 μm and 1.8-1.96 μm) were removed.

species identities in sufficiently detailed optical data. The methods that decompose these signals to map species traits and identities, or patterns of ecosystem structure and function, are referred

to throughout this work as methods for *measuring* biodiversity patterns. Due to multiple forms of covariance—including covariance between traits as well as covariance from each trait propagating signal across the reflectance spectrum—disentangling the effects of each factor is a major challenge. *Chapter 2* is a review of some of the scale-dependent challenges to signal decomposition in the context of biodiversity mapping, and *Chapter 3* describes a machine learning method that overcomes some of these challenges through ecologically-informed data preprocessing to identify tree species in high resolution imagery.

In addition to passive optical sensors, there is a diverse array of earth observing sensors that measure other biologically and ecologically important patterns. Thermal sensors measure midwave (3-8 μm) and longwave (8-10 μm) emmissivity patterns, which can be used to map surface temperature patterns [Li et al., 2013]. Microwave instruments are sensitive to the volume of liquid or atmospheric water, depending on the sensor, tracking patterns ranging from precipitation to canopy water content [Hou et al., 2013, Konings et al., 2019]. Even cloudy pixels in optical data, which make up around 60% of observations and are generally a nuisance for biodiversity mapping, have been used to refine the extents of cloud forests and improve predictions of fog-dependent species distributions [Wilson and Jetz, 2016].

These patterns are useful in biodiversity monitoring contexts despite violating the "biological" requirement to be considered Essential Biodiversity Variables. Niche use patterns are often defined by temperature, precipitation, and cloud cover patterns, placing abiotic constraints on the spatial distributions of species (which are indeed biological, sensitive to change, and can be mapped across ecosystems). Since these measurements are often used as predictive features in spatial biodiversity models, these data are referred to throughout this document as useful for *modeling* biodiversity patterns. *Chapter 2* explores the scale-dependent challenges of modeling biodiversity patterns with earth observations data, and is especially concerned with species distribution modeling examples. *Chapter 4* applies the lessons learned from those use cases to modeling the spatial distributions and niche use patterns of two globally-invasive mosquito arbovirus vectors, *Aedes aegypti* and *Ae. albopictus*, which are expected to continue expanding to new regions as temperature and precipitation patterns shift with climate change [Tjaden et al., 2018, Ryan et al., 2019].

1.3   Dissertation Themes

There are three general themes explored in this dissertation. First is regarding the role of spatial and temporal measurement scales in determining which biodiversity patterns can be detected by earth observations sensors, including what drives signal variance at each scale. Changing measurement scales, both by definition and in practice, changes the biological variation within and between measurements, and nonlinear patterns often emerge at ecological boundaries known as domains of scale. The transition from measuring leaf reflectance to measuring canopy reflectance is an example of these nonlinear effects. Light traveling through a canopy is absorbed, transmitted and reflected by individual leaves, and the transmitted and scattered photons can be absorbed, transmitted, or reflected by other leaves within the canopy. Since the amount of light absorbed, transmitted, or reflected varies by wavelength, plant canopies are spectrally distinct from the individual leaves that comprise them, with dense canopies absorbing more visible light and scattering and reflecting more infrared light. This is the mechanistic basis for retrieving forest structural patterns like leaf area index from optical data [Jordan, 1969]. While some changes propagate across scales, like changes in chlorophyll concentrations, the methods for measuring these changes depends on the scale of analysis, as canopy structure introduces additional spectral variance that covaries with pigment changes. Identifying the appropriate domains of scale for analysis is a critical when measuring and modeling biodiversity patterns, and the example above is explored in-depth in *Chapter 3*.

The second theme of this dissertation concerns the Goldilocks principle. Considering the global nature of the field, ecology was previously a data-scarce discipline, as many central ecological theories were derived from data on a single species, a single community, or a few islands [Grinnell, 1917, Hubbell, 2001, MacArthur and Wilson, 1967]. The past decade has upended this trend, as open and standardized biodiversity data are more abundant and accessible than ever. So, too, are global earth observations data. But this flood of information can be overwhelming. This dissertation explores approaches to combining incomplete species data ("too little") with abundant earth observations data ("too much") to map patterns of biodiversity that are "just right." *Chapter 3* decomposes 426 spectral reflectance bands from airborne imaging spectroscopy data into principal components finding that, with feature selection, only around 20-30 components were required to discriminate between species. *Chapter 4* analyzes over 8,000 species occurrence records and distills 16 years of daily satellite imagery into descriptive metrics of climate, habitat and resource use across Latin

America and the Caribbean to quantify constraints on niche use for *Ae. aegypti* and *Ae. albopictus*, mapping the potential distributions of these arbovirus disease vectors at continental scales.

The final theme of this dissertation is the use of biomimicry: selecting and training machine learning models to approximate mechanistic ecological models. Commensurate with the expansion of open data, machine learning methods for analyzing ecological data have recently proliferated [Pal, 2005, Phillips et al., 2006, Hastie et al., 2009, Elith et al., 2011, Brodrick et al., 2019]. These are essential tools for identifying novel patterns in large datasets where mechanistic relationships are unclear. But they are also prone to overfit to spurious, non-biological patterns, especially in spatial contexts [Hawkins, 2012, Fourcade et al., 2018]. I approach the problem in this dissertation by carefully selecting machine learning models that mimic well-known ecological processes, using data that has been transformed to fit model assumptions. *Chapter 3* uses decision tree ensembles to predict tree species identities based on decomposed reflectance data, approximating a dichotomous tree and morphological trait-based taxonomy. *Chapter 4* fits non-linear and smoothly varying response functions to climate, habitat, and resource use patterns for *Ae. aegypti* and *Ae. albopictus*, which was designed to mimic thermal response functions characterized for those vectors in laboratory settings [Mordecai et al., 2019].

Together, this work further develops the technical and conceptual basis for building biodiversity monitoring systems that carefully integrate earth observations and *in situ* data. However, much work remains to actually build these systems and provide the data to the right stakeholders who can use it to inform conservation and land management decisions [Guerry et al., 2015, Ramirez-Reyes et al., 2019]. The end of the last decade marked the time to deliver on the promises of the Ecosystem Services framework; now is the time to deliver on the promise of biodiversity monitoring [Daily et al., 2009, IPBES, 2019].

1.4 Publication and Copyright Notices

Two of the following chapters have been published in peer-reviewed journals under open-access licenses. Chapter 2 was published in *Ecology Letters* in July 2018, which is reprinted with permission from the publisher, John Wiley and Sons Ltd. The full license agreement is included in *Appendix D*. The full citation is:

> Anderson, C. B. (2018). Biodiversity monitoring, earth observations and the ecology of scale. *Ecology letters*, 21(10), 1572-1585.

Chapter 3 was published in *PeerJ* in October 2018 with the following full citation:

> Anderson, C. B. (2018). The CCB-ID approach to tree species mapping with airborne imaging spectroscopy. *PeerJ*, 6, e5666.

Per the *PeerJ* Copyright Policy (`https://peerj.com/about/policies-and-procedures/`):

> All *PeerJ* articles are published under a Creative Commons Attribution License. With this license, Authors retain copyright, but allow any user to share, copy, distribute, transmit, adapt and make commercial use of the work without needing to provide additional permission, provided appropriate attribution is made to the original author or source.

Chapters 1, 4, and 5 are currently unpublished outside of this dissertation.

Chapter 2

Pattern and Scale in Ecology and Earth Observations Science

Christopher B. Anderson

2.1 Abstract

Human activity and land-use change are dramatically altering the sizes, geographical distributions and functioning of biological populations worldwide, with tremendous consequences for human well-being. Yet our ability to measure, monitor and forecast biodiversity change–crucial to addressing it—remains limited. Biodiversity monitoring systems are being developed to improve this capacity by deriving metrics of change from an array of *in situ* data (e.g. field plots or species occurrence records) and Earth observations (EO; e.g. satellite or airborne imagery). However, there are few ecologically based frameworks for integrating these data into meaningful metrics of biodiversity change. Here, I describe how concepts of pattern and scale in ecology could be used to design such a framework. I review three core topics: the role of scale in measuring and modelling biodiversity patterns with EO, scale-dependent challenges linking *in situ* and EO data and opportunities to apply concepts of pattern and scale to EO to improve biodiversity mapping. From this analysis emerges an actionable approach for measuring, monitoring and forecasting biodiversity change, highlighting opportunities to establish EO as the backbone of global-scale, science-driven conservation.

2.2 Introduction

Global biodiversity monitoring is a crucial but challenging task, as human activities are changing the structure and composition of biological populations at all taxonomic levels [Dirzo et al., 2014, Ceballos et al., 2017]. Mitigating biodiversity loss will require understanding the rates, magnitudes and geography of these changes [Laurance et al., 2012, Mendenhall et al., 2014]. However, considering the scope of action required for mitigation, our knowledge of global biodiversity change remains limited [Daily, 1999, Pereira et al., 2012]. Furthermore, what is known about biodiversity change is complicated by taxonomic, geographical and recency biases [Boakes et al., 2010, Donaldson et al., 2016, Gonzalez et al., 2016].

Novel biodiversity monitoring systems are being developed to overcome these biases by systematically assessing change for multiple taxa over large extents [Scholes et al., 2008, Scholes et al., 2012, Fernández et al., 2015]. To support these systems, several groups have developed novel approaches to monitor species, communities and ecosystems over time using globally consistent metrics of change [Butchart et al., 2010, Jetz et al., 2012, Metzger et al., 2013, Pereira et al., 2013]. These metrics are biological, sensitive to change, and ecosystem agnostic, enabling consistent monitoring protocols worldwide [GEO BON, 2017]. These efforts have been greatly bolstered by increasing access to globally available *in situ* biodiversity observations [Geijzendorffer et al., 2016, Culina et al., 2018]. However, as *in situ* data alone are often insufficient for assessing global diversity patterns (*sensu* the Linnean and Wallacean shortfalls [Bini et al., 2006, Brito, 2010]), researchers have looked for complementary data to support monitoring efforts.

Earth observations (EO; e.g. satellite or airborne imagery) complement *in situ* data by providing repeat, thematically consistent and spatially continuous measurements of terrestrial ecosystems, characterising biodiversity patterns over large, undersampled areas. However, linking fine-scale field and EO data faces many challenges, including incomplete sampling efforts (i.e., where field measurements do not adequately characterise the extent of environmental variation; [Marvin et al., 2014]) and reconciling scale mismatches (e.g., where field plots are much smaller than EO measurements). Developing EO-based biodiversity monitoring systems will require a comprehensive approach to link these data [Turner, 2014, Pettorelli et al., 2016].

Scale plays a key role in both ecology and EO science, and identifying shared scaling dynamics could provide a basis for bridging these disciplines. Understanding the roles of spatial and temporal

scales in biological communities is a central topic in ecology, and is referred to as the problem of pattern and scale [Wiens, 1989, Levin, 1992]. The problem of pattern and scale emphasises that multiple ecological processes often determine the spatial distributions of biodiversity patterns, and that these processes can act across multiple spatial and organismal scales [Withers and Meentemeyer, 1999, Waring and Running, 2010, Chase and Knight, 2013]. Therefore, there is rarely a single measurement scale that best identifies how specific processes drive patterns [Hutchinson, 1953]. EO measurements are subject to similar scale dependencies: the grain size of an EO sensor often determines which patterns can be measured (Fig. 2.1) [Lechner, 2010, Anderson, 2012, Nagendra et al., 2013], and multi-scale EO analyses can reveal the influences of multiple processes driving biodiversity patterns [Keil et al., 2012, Taylor et al., 2015]. Applying concepts of pattern and scale in ecology to EO could provide a means to better link these fields, paving the way for improved biodiversity monitoring.

Here, I review the scales at which EO have been used to measure and model metrics of biodiversity change, and the role of scale in linking field data with EO. This is not strictly a review of which biodiversity patterns EO can measure (*sensu* [Roughgarden et al., 1991, Turner et al., 2003, Wang et al., 2010, Pettorelli et al., 2014a, Lausch et al., 2016]). Instead, this review addresses three questions:

1. At what scales have current and historical EO been used to measure or model spatial biodiversity patterns?

2. What are the major challenges linking field-based and EO-based biodiversity measurements, and how does scale impact these challenges?

3. How can concepts of pattern and scale, applied to EO, facilitate the translation of biodiversity patterns across scales?

2.3   Components of Pattern and Scale in Ecology

Explorations of pattern and scale in ecology focus on two distinct but related measurement scales: grain size and extent. In this review, I refer to these scales in a spatial sense, though temporal grain size could describe the frequency of observations (e.g., one diurnal cycle for net primary productivity) and temporal extent could describe the total time over which an ecological process
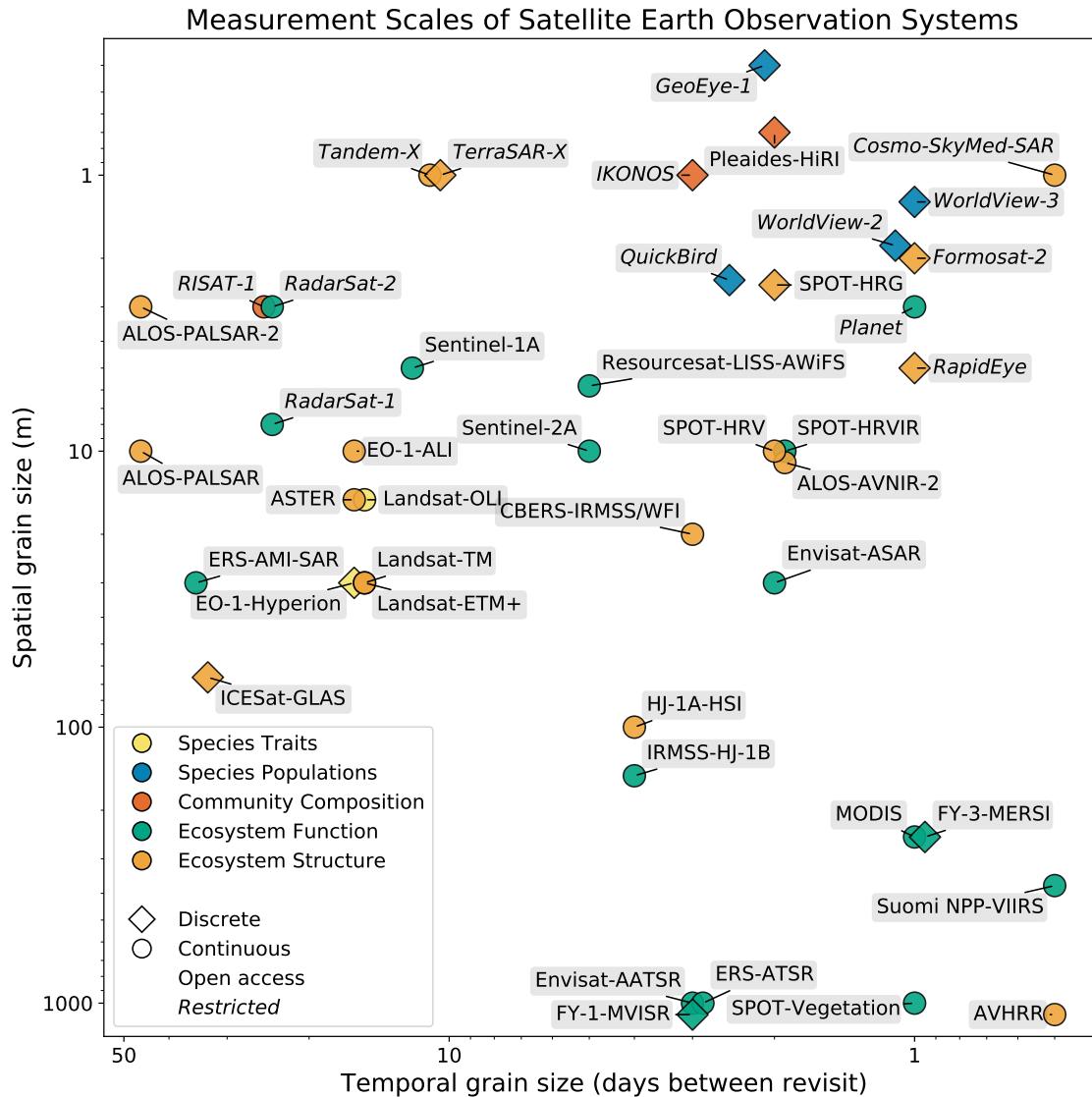
Figure 2.1: Spatial and temporal and grain sizes for 44 current and historic satellite Earth observation (EO) sensors, colored by EBV class. Several sensors have been used to measure multiple biodiversity patterns, and the most cited or most novel were selected in these cases.

occurs (e.g. phenological variation throughout a year). Furthermore, I adopt the classes and metrics of biodiversity change from the Essential Biodiversity Variables framework [Pereira et al., 2013], and refer to these metrics as biodiversity patterns. This framework captures the multiple biological scales of diversity (i.e., variation in genes, species, communities and ecosystems) as opposed to a

more narrow interpretation that refers to biodiversity as variations in species richness, abundance and evenness. I believe these disaggregated classes and metrics more comprehensively address the patterns that can be measured and modelled using EO. In this section I discuss how concepts of pattern and scale in ecology apply in biodiversity and EO contexts, then I review domains of scale, which constrain efforts to generalise patterns across scales. See the *Glossary* for clarifications on the terms and acronyms used.

2.3.1 Changing Measurement Scales

Measurement scales are often selected to understand biodiversity patterns or ecological processes in a specific region, or for a specific species, population or ecosystem. A key scaling dynamic is that when the scale of measurement changes, the variation within that measurement is also subject to change [Wiens, 1989, Levin, 1992]. For example, early biodiversity/ecosystem function research suggested the relationship between species richness and productivity to be unimodal, predicting peak biomass accumulation at intermediate diversity for both primary and secondary productivity [Rosenzweig and Abramsky, 1993]. However, this functional form was shown to be an artefact of plot size as opposed to any ecological process [Oksanen, 1996], and a global synthesis found mixed evidence for a generalised relationship in any form [Adler et al., 2011]. Recently, however, long-term studies addressing scale directly have demonstrated a positive yet saturating diversity-productivity relationship across multiple ecosystem [Liang et al., 2016, Hungate et al., 2017].

Measurements of community-scale patterns, like species richness and turnover (i.e. α and β diversity), have also been shown to vary directly with scale [Rosenzweig, 1995]. Coarse grains are expected to contain higher species richness per grain, and thus lower species turnover between grains [Nekola and White, 1999, Whittaker et al., 2001]. This is because larger grains are expected to contain more rare species and more environmental variation (e.g. more variation in niche space; [Keil et al., 2015]. Indeed, [Hurlbert and Jetz, 2007] showed systematic increases in species richness at coarser grain sizes for birds in South Africa and Australia. Similarly, species turnover has been shown to decrease at coarser grains for birds in Britain and North America [Mac Nally et al., 2004, Gaston et al., 2007], and for mammals in Mexico [Arita and Rodriguez, 2002].

Measurement scales likewise determine which biodiversity patterns can be measured by EO (Fig. 2.1). Satellite EO have historically focused on measuring ecosystem-scale patterns, due to the

coarse grain sizes of historic sensors. Coarse grain EO sensors measure ecosystem-scale patterns, like disturbance regime [Wang et al., 2012, Kogan et al., 2015] and ecosystem extent [Maillard et al., 2008, Bartsch et al., 2009]. Fine-grain sensors measure species- and community-scale patterns like species occurrences [Immitzer et al., 2012] and taxonomic diversity [Khare et al., 2018]. Measuring species traits has likewise proven challenging due to difficulties distinguishing individual organisms in EO imagery [Nagendra et al., 2013, Jetz et al., 2016]. But some plant traits, like canopy nitrogen content and photosynthetic rates, can be retrieved even at moderate grain sizes [Martin et al., 2008, Serbin et al., 2014]. High frequency measurements can map temporally sensitive processes like vegetation phenology [Bradley et al., 2007], but high frequency, continuous measurements often come at the expense of coarser grain sizes. Fortunately, an increasing number of fine-grain EO sensors now in orbit is enabling EO biodiversity mappers to focus on more species- and community-scale patterns (Fig. 2.2) [Butler, 2014].

There are key similarities in scaling dynamics between field and EO data: grain size and extent both constrain within and between-grain measurement variation. Large field plots tend to contain more species per plot, and lower turnover between plots. Likewise, large EO pixels tend to contain more organisms per grain, and lower turnover between grains. This constrains measurement specificity. However, the grain size of an EO sensor only constrains the smallest unit of measurement; these data can be spatially aggregated to larger scales [Fisher, 1997]. For example contiguous pixels measuring the same tree could be aggregated to delineate a single crown, or clusters of forested pixels could be aggregated to delineate forest fragments [Yao et al., 2015]. This enables comparisons between crowns or across fragments, instead of pixels, helping bridge the gap between spatial and biological scales. This is known as object-based image analysis [Blaschke et al., 2008], which is likely to become more common in biodiversity monitoring as novel segmentation algorithms are tuned for EO [Krizhevsky et al., 2012, Basu et al., 2015]. And though this approach facilitates ecological interpretations of EO data, there are key scaling dynamics associated with aggregating data across scales.

2.3.2   Domains of Scale

One tenet of the problem of pattern and scale in ecology suggests that, since multiple ecological processes often drive spatial biodiversity patterns, there is rarely a single scale at which any pattern must be examined [Hutchinson, 1953, Levin, 1992]. These patterns are often examined at multiple

points along biological, spatial or temporal scale spectrums in order to understand how multiple processes drive patterns. For example, the drivers of net primary productivity in plants could be examined at leaf, whole plant and landscape scales. The leaf, plant and landscape, here, represent domains of scale: the scales over which patterns either do not change, or change monotonically with changes in scale [Wiens, 1989]. In this example, fine-scale processes, like intra-crown shading, may drive the majority of variation in leaf-scale productivity, but may be less important at landscape scales, where ecosystem processes like resource availability drive the majority of variation [Field et al., 1995]. Partitioning biodiversity patterns into genetic, species, community and ecosystem-scale patterns organises them as domains of scale; the processes that drive variation in species-scale patterns are expected to drive variation in ecosystem-scale patterns through separate but potentially nested pathways [Pereira et al., 2012, Pereira et al., 2013].

Constraining measurements and models to discrete domains of scale is key for simplifying predictions of how species respond to environmental change [Field, 1991]. Multi-scale analyses have been used to identify domains of scale, revealing where transitions across scales has nonlinear effects on observed patterns [Palmer and White, 1994]. In community ecology, hierarchical regression models have been employed to this end [Legendre et al., 2005]. For example, [Keil et al., 2012] tested how β-diversity patterns for birds, butterflies, plants, amphibians and reptiles across Europe varied with distance, climate and land cover. They found β-diversity (here, dissimilarity) decreased systematically at coarser grain sizes for each taxon. Their hierarchical analysis found climate was important for predicting β-diversity patterns at coarse grain sizes, and land cover was important at fine-grain sizes, though these effects varied by taxon. Their results suggest that efforts to predict changes in turnover should assess the effects of multiple domains of scale simultaneously, and that these scale dependencies are taxon-specific.

The domains of scale where processes drive patterns may not always be known *a priori*, however. These are often identified using multi-scale sensitivity analysis. For example, [Mendenhall et al., 2011] developed a multi-scale model to predict how bird community composition changed with land cover change in Costa Rica. They assessed species turnover along tree cover gradients, finding turnover varied nonlinearly with cover at both fine and coarse grain sizes. Their results suggested there are two domains of scale over which tree cover patterns determine turnover patterns for birds (perhaps tracking habitat and resource availability, respectively; [Morrison et al., 2012]). Furthermore, their
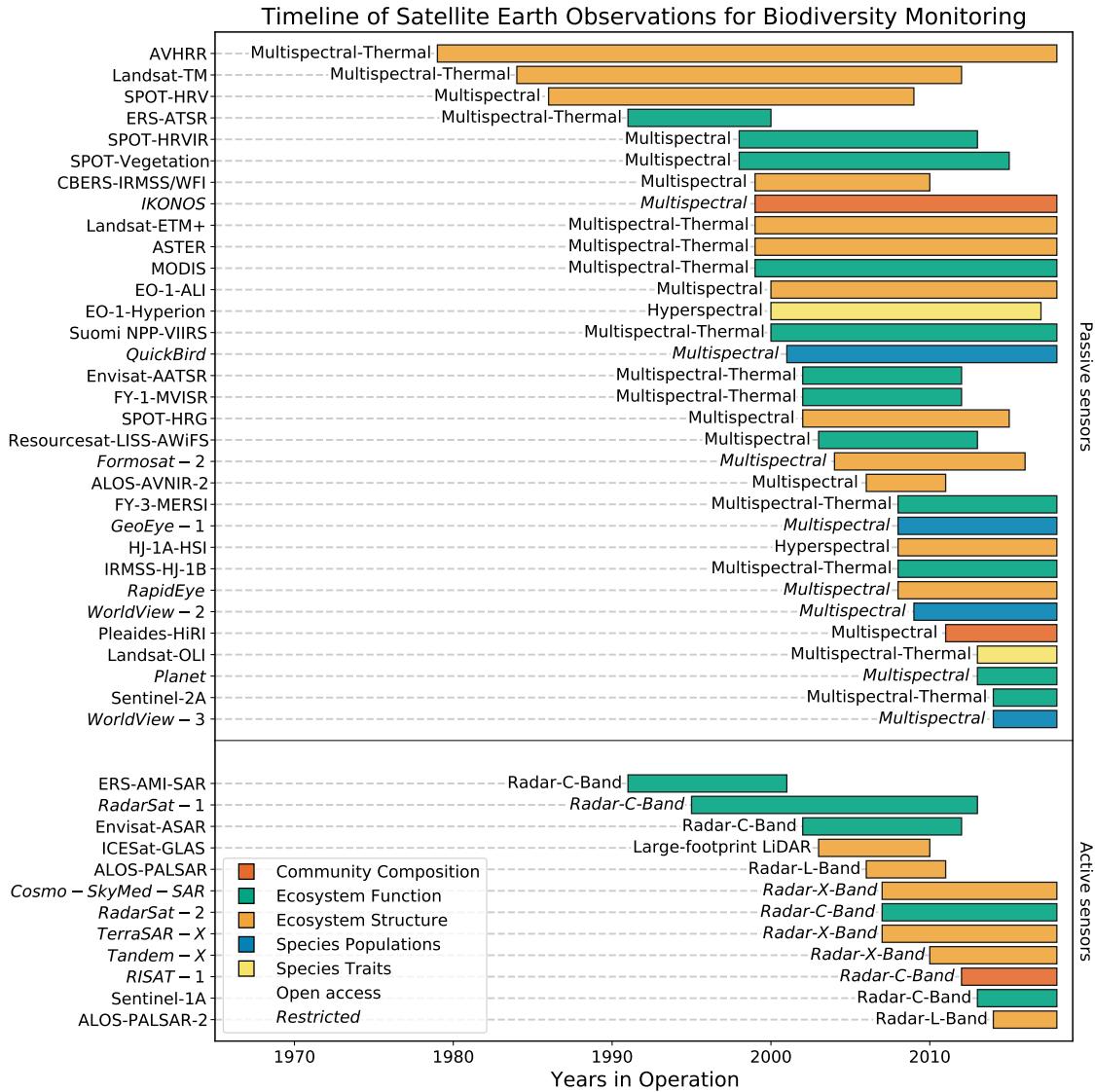
Figure 2.2: Temporal extents of current and historic satellite sensors, colored by EBV class. Temporally coincident measurements can be jointly analyzed in a multi-sensor fusion framework to increase data dimensionality.

results suggested tree cover change could serve as a proxy to predict turnover in other communities. Indeed, [Mendenhall et al., 2016] found tree cover change predicted changes in composition for understory plants, non-flying mammals, bats, reptiles and amphibians. Furthermore, they found the grain size of tree cover which best predicted turnover varied by taxon. Their work highlights one

approach to mapping biodiversity change with EO – identifying domains of scale through multi-scale sensitivity analysis, then modelling turnover via regression with EO-derived environmental features.

2.4 Measuring and Modeling Biodiversity Patterns with EO

There are currently two principal paradigms for mapping biodiversity patterns with EO [Turner et al., 2003]. First is to directly measure species, community or ecosystem-scale patterns. Examples of this paradigm include identifying individual organisms within a species [Gairola et al., 2013] or mapping the extent of an ecosystem [Henderson and Lewis, 2008]. Second is to model biodiversity patterns indirectly using EO as predictive environmental features. Examples of this paradigm include modelling species richness from measurements of habitat structure [Saatchi et al., 2008], or modelling species distributions and turnover using land cover maps [Guisan and Thuiller, 2005, Keil et al., 2012]. Here I review the roles of measurement type and measurement scales in these paradigms, focusing on biodiversity patterns mapped by current and historic spaceborne sensors that can be accessed by biodiversity monitoring systems (Table B.1).

2.4.1 Measuring Biodiversity Patterns

EO measurements of biodiversity patterns are characterised by three key properties: sensor type, sensor fidelity and measurement scales [Pettorelli et al., 2014b, O'Connor et al., 2015]. Sensor type determines which patterns can be measured, sensor fidelity constrains the variation in those measurements, and measurement scales determine the amount of variation within and between measurements [Jensen and Lulla, 1987]. Passive sensors, such as multispectral sensors and imaging spectrometers, measure patterns of ecosystem function, like leaf area index [Fensholt et al., 2004], vegetation phenology [Fan et al., 2015] or disturbance regime [Feng et al., 2008]. Active sensors, such as radio or light detection and ranging sensors (i.e., radar and lidar), often measure patterns of ecosystem structure, like tree height [Lefsky et al., 2005] and ecosystem extent [Bartsch et al., 2009]. These distinctions are not axiomatic; multiple sensor types have been used to measure the same pattern [Pohl and Van Genderen, 1998]. Both radar and multispectral sensors have been used to measure tree cover, for example. Radar sensors map tree cover by measuring woody structural and hydrological characteristics [Walker et al., 2010, Shimada et al., 2014], and multispectral sensors map tree cover by measuring leaf optical properties like pigment concentrations [Sims and Gamon,

2002, Sexton et al., 2013].

Using multiple sensors to map a single biodiversity pattern can improve model accuracy and reduce sensor-specific uncertainties, and is known as multi-sensor fusion [Hall and Llinas, 1997]. Consider, again, the application for tree cover mapping. Though multispectral sensors are sensitive to pigment concentrations, measuring tree cover in leaf-off conditions is a challenge; exposed branches are optically similar to dried grass or other non-photosynthetic vegetation [Asner, 1998]. Radar is sensitive to woody biomass regardless of phenology. but can itself be noisy due to speckling [Lee et al., 1994]. To obviate this issue, [Naidoo et al., 2016] mapped cover in a South African savannah by combining multispectral and radar measurements. Combined, they mapped tree cover with 90% accuracy, which was 12% higher than using either sensor independently. Multi-sensor fusion approaches to biodiversity mapping hold great promise for reducing sensor-specific uncertainties, and are poised to become more valuable as access to novel sensor types increases (Fig. 2.2) [Butler, 2014, Schulte to Bühne and Pettorelli, 2018].

Comparing measurements from similar sensor types with different grain sizes has been used to assess the importance of scale in measuring biodiversity patterns. For example [Brown et al., 2006] compared NDVI measurements from four spaceborne multispectral sensors and found that up to 20% of the measurement variance between sensors was driven by differences in grain size. Furthermore, [Garrigues et al., 2006] found that changes in grain size explained to up to 50% of the variance in comparisons of multi-scale leaf area index measurements, which increased at coarser grains and in spatially heterogeneous landscapes. Comparing these spatial uncertainties to the radiometric calibration uncertainties of EO sensors (i.e. sensor fidelity), which are often between $\pm 5 - 10\%$ absolute radiance [Chander et al., 2009], suggests that differences in measurement scales can be at least as important as differences in sensor fidelity for measuring biodiversity patterns. The physical drivers of this scale dependence have been explored with radiative transfer models, particularly for patterns of ecosystem function [Asner et al., 1998, Jacquemoud et al., 2009], but should be further quantified for other biodiversity patterns.

2.4.2   Modeling Biodiversity Patterns

Biodiversity patterns that are difficult to measure directly with EO are often modelled as a function of environmental features (Fig. 2.3). There are many approaches to modelling biodiversity patterns
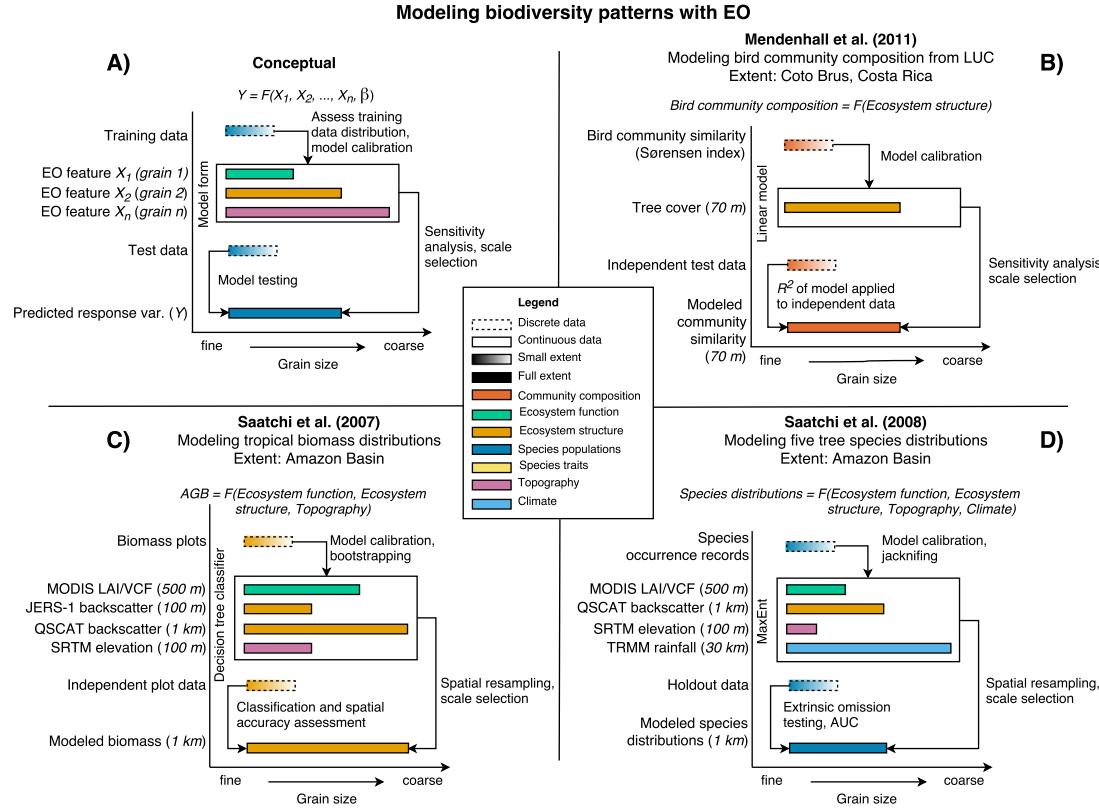
Figure 2.3: Biodiversity patterns are often modelled using EO data at a single spatial scale. (a) Conceptual model of how EO have been used to predict biodiversity patterns using supervised modelling approaches. (b) [Mendenhall et al., 2011] modeled bird community similarity as a function of tree cover across Coto Brus, Costa Rica. (c) [Saatchi et al., 2007] modeled aboveground biomass distributions across Amazonia. (d) [Saatchi et al., 2008] modeled tree species distributions across Amazonia.

with EO, including models of species-scale (Fig. 2.3d) [Saatchi et al., 2008], community-scale (Fig. 2.3b) [Mendenhall et al., 2011] and ecosystem-scale biodiversity patterns (Fig. 2.3c) [Saatchi et al., 2007]. These approaches typically resample all data layers to a uniform grain size and extent, occasionally after multi-scale sensitivity analysis [McGarigal et al., 2016]. Here I briefly discuss models of individual species distributions [Guisan and Thuiller, 2005] and models of community-scale patterns like α- and β-diversity [Rocchini, 2007]. These modelling methods have been reviewed elsewhere [Gillespie et al., 2008, Rocchini et al., 2010, Pettorelli et al., 2014b], but this section reviews role of scale in these approaches.

Species distribution models (SDMs) predict species geographical distributions across an extent as a function of environmental features that constrain niche use and niche availability [Soberon and Peterson, 2005]. There have been many discussions on the importance of feature selection in SDM [Booth et al., 2014, Brandt et al., 2017, Fourcade et al., 2018], but some key reviews have emphasised that scale selection can play a similarly important role [Mayor et al., 2009, McGarigal et al., 2016]. Even so, studies addressing scale directly have found equivocal results. For example [Guisan et al., 2007] modelled bird and plant distributions at multiple grain sizes, finding only small decreases in model accuracy at coarser grain sizes on average. Disaggregating these results by taxon, however, revealed significant decreases in accuracy at coarser grains for all plants, but only some birds. In addition, species with the least training data saw the largest decreases in accuracy. [Seo et al., 2009] further explored these patterns in nine plant species, comparing both model accuracy and the spatial patterns of distributions. They found model accuracy decreased consistently at coarser grain sizes, and that these decreases were species-specific. They also found significant spatial disagreement between models of varying grain size for each species, which could have major consequences for spatial conservation planning [Faleiro et al., 2013].

There are two principal approaches to modelling community-scale patterns with EO. First is to predict the distributions of all species in a community, then overlay these outputs to estimate community composition (i.e. stacked SDMs; [Thuiller et al., 2009, Calabrese et al., 2014]. Second is to model community diversity metrics via regression [Gillespie et al., 2008, Saatchi et al., 2008]. As above, the role of scale in these approaches has been equivocal. For example [Thuiller et al., 2015] modelled multiple plant community diversity metrics in the French Alps using a stacked-SDM approach at varying grain sizes. They found that estimates of functional diversity, phylogenetic diversity and species richness all varied independently with changes in grain size. Functional diversity was best predicted at the finest grain size (250 m), whereas phylogenetic diversity and species richness were best predicted at coarser grain sizes (1000 m), suggesting scale dependence at the community scale is often process specific.

Assessing scale dependence in regression approaches has been done by comparing species richness predictions across multiple sensors. For example [Nagendra et al., 2010] modelled plant species richness using features from a fine grain, low fidelity sensor (IKONOS) and a moderate grain, high fidelity sensor (Landsat). Since community diversity metrics assess within- and between-grain

variation, one may expect that fine-grain EO better predict these patterns. On the other hand, high fidelity measurements may better discriminate the between-grain variation in environmental features that predict spatial turnover in communities. [Nagendra et al., 2010] found that, despite the coarser grain size, Landsat-based models better predicted plot-level species richness. Though the IKONOS data matched the grain size of the field plots, they failed to meaningfully discriminate the spatial variation in environmental features that predicted spatial richness patterns. Further disentangling the effects of sensor fidelity from variations in measurement scales will help discriminate sensor dependence from scale dependence in modelling other biodiversity patterns.

2.4.3 Linking Measurements and Models

EO measurements and models of biodiversity patterns are tightly connected. They are both subject to pattern-specific scale dependencies, and multi-scale comparisons or sensitivity analyses are essential for quantifying and understanding these dependencies. Furthermore, when EO measurements are the features used to model biodiversity patterns, scale-dependent measurement variation becomes embedded within the models. This might obfuscate process-driven scale dependence for variation driven by changing measurement scales. Constraining scale-dependent variation in EO measurements of biodiversity patterns, and disentangling this variation from variation driven by sensor fidelity, will be key for reducing uncertainties in multi-scale modelling efforts. In the following section I review some other challenges linking measurements and models of biodiversity patterns, and opportunities for multi-scale analyses to address these challenges.

2.5 Translating Biodiversity Patterns Across Scales

One central challenge linking field and EO data is overcoming scale mismatches. These mismatches occur where response and feature data are sampled at disparate and irreconcilable scales. The size of field plots (i.e., the response data) are often much smaller than the grain size of EO sensors (i.e., the feature data), which can obscure key patterns and processes operating between these scales. For example [Cleveland et al., 2015] modelled spatial patterns of net primary productivity across the Amazon basin using three models at three scales: from plot data upscaled to the study extent ($0.1\,ha$ grain size), from MODIS data collected across the full extent ($1\,km^2$ grain size) and from a community land model ($12,500\,km^2$ grain size). These methods calculated the same average

net primary productivity across the Amazon, indicating a potential convergence of the processes driving forest productivity. However, results from the finer-scale methods were shown to be spatially independent from the others, indicating that they converged on the same average results for different reasons. In this case, comparing multiple models at mismatched scales that calculated the same result can make it difficult to disaggregate the role of process from the role of scale in understanding spatial patterns of productivity.

One key challenge in translating patterns across mismatched scales is capturing the dynamics of intermediate-scale biodiversity patterns that are poorly characterised by field data. These patterns are too rare to be characterised by a small number of field plots, and are often difficult to reliably measure with coarse grained EO. For example, [Fisher et al., 2008] and [Chambers et al., 2009] identified that tree falls patterns, which tend to be both rare and spatially clustered, are underrepresented in field plots in the Amazon. Their analyses demonstrated that efforts to model related patterns using just field data (e.g., carbon sequestration) would necessarily underestimate feedbacks from these intermediate-scale disturbances. Furthermore, [Marvin et al., 2014] quantified these mismatches using airborne lidar data, finding between 44 and 85 field plots per forest type would be required to characterise mean, community-scale carbon and disturbance dynamics. Together, these results suggest that field measurements should be greatly expanded, or novel data should be used to characterise intermediate-scale patterns in order to translate patterns across scales.

The challenges presented by scale mismatches can be framed by two tenets of the problem of pattern and scale: that multiple ecological processes can drive biodiversity patterns, and that there is rarely a single scale that best identifies how specific processes drive patterns [Wiens, 1989, Levin, 1992]. These tenets suggest that multi-scale analyses, which capture the intermediate-scale patterns obscured between fine and coarse grain patterns, could improve empirical approaches to mapping biodiversity patterns with EO. Iteratively modelling patterns with multi-scale EO has been used to map a range of biodiversity patterns at moderate grain sizes across large extents, and presents an actionable approach to overcoming some of the challenges presented by scale mismatches.

2.5.1 Multi-scale Modeling

Multi-scale models attempt to obviate scale mismatches through iteratively modelling patterns at varying grain sizes (Fig. 2.4a). One key innovation of the multi-scale modelling approach was to
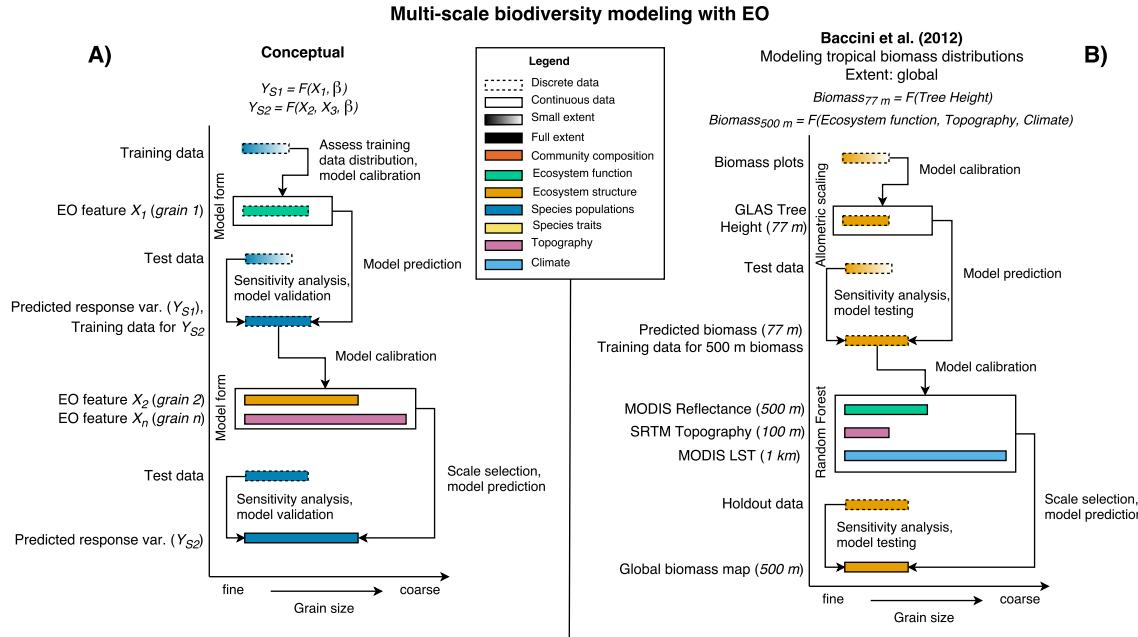
Figure 2.4: Leveraging coincident data from multiple sensors in a multi-scale framework provides opportunities to translate fine-scale biodiversity patterns across scales. (a) Conceptual model of how EO data have been used to predict biodiversity patterns using a multi-scale modelling approach. (b) An example from [Baccini et al., 2012] demonstrating the multi-scale modelling approach for predicting tropical biomass globally.

leverage intermediate-scale data sources that capture the extent-wide variation in EO features, which is difficult to cover with field plots alone. For example, [Baccini et al., 2012] developed a benchmark map of pantropical aboveground biomass using a multi-scale model, a network of field plots, discrete spaceborne lidar data $(4,900\,m^2)$ and continuous, coarse grain EO $(0.25\,km^2$; Fig. 2.4b). First, they calibrated an allometric model using biomass plots coincident with lidar-derived tree height data *sensu* [Chave et al., 2005]. Next, they applied this model to all tree height measurements, creating a discrete, global biomass map. Finally, they modelled biomass continuously using a regression tree model, with lidar-derived biomass as the response and EO data on climate, topography and ecosystem function as the environmental features. Their final map of aboveground biomass served as a benchmark for global carbon monitoring [Ciais et al., 2014].

Multi-scale models have also been used to monitor temporal changes using intermediate-scale EO, overcoming some of the challenges highlighted by [Fisher et al., 2008] and [Marvin et al., 2014]. For

example, [Baccini et al., 2017] assessed temporal patterns of change in aboveground biomass using EO measurements of forest growth, disturbance and deforestation. Intermediate-scale disturbance measurements were essential for capturing the magnitude of change: their results revealed that disturbance accounts for nearly 70% of forest emissions, and that the Earth's tropical forests are now a net source of carbon to the atmosphere. Considering how little is known about the rate, magnitude and direction of global biodiversity change [Pereira et al., 2012, McGill et al., 2015], I expect these multi-scale analyses will prove essential for settling debates over other key knowledge gaps [Vellend et al., 2013, Gonzalez et al., 2016]. These analyses have also proven useful for mapping patterns that have been difficult to directly measure with publicly-accessible EO data: species traits.

2.5.2 Mapping Plant Functional Traits

Measuring species traits as a complement to species counts has become a priority for biodiversity science. Traits have been touted as a link between applied and theoretical biodiversity research and as a means to better represent ecosystem function in Earth systems models [Shipley et al., 2006, Jetz et al., 2016, Funk et al., 2017]. Plant functional traits (PFTs) are one subset of species traits that can be mapped using EO, specifically by imaging spectroscopy [Kokaly et al., 2009]. One key benefit of measuring PFTs with EO is that they can be mapped without having to identify and characterise every species *a priori*; capturing the range and variation in traits is often more important. The prospective launches of spaceborne imaging spectrometers, such as EnMAP, PRISMA and HISUI, are currently touted as the best bet for mapping PFTs globally [Stuffler et al., 2007, Galeazzi et al., 2008, Matsunaga et al., 2011]. Simulations from preparatory campaigns have found mixed results, however. For example, [Bachmann et al., 2015] demonstrated that the moderate fidelity of these sensors should lead to high variation in surface reflectance measurements (the basis for measuring PFTs). Furthermore, the moderate grain size of these sensors (30 m) has been shown to significantly reduce classification accuracy compared to fine-grain measurements in other contexts [Kruse et al., 2011]. This decrease in accuracy is expected to be exacerbated for PFTs since canopy structure, not trait variation, drives the majority of reflectance signal at moderate grain sizes [Yao et al., 2015]. Combined, these sensor- and scale-dependent challenges spell major trouble for trait-based species mapping. This approach is predicated on the basis that differences in traits between species arise as a function of interpsecific niche partitioning and intraspecific niche conservation. That is, traits are

typically conserved within a species and vary between species along axes of e.g. the plant economics spectrum [Wright et al., 2004]. Tracking these patterns with EO could be for naught if intraspecific measurement variance is greater than the variance in interspecific measurements.

The implication is that spaceborne trait measurements may not yet provide a panacea for large-scale species mapping. Fortunately, airborne imaging spectrometers can measure species traits at the scales of individual organisms, and these measurements can be combined with other EO to model PFT distributions over large extents (Fig. 2.4a), as [Asner et al., 2016] demonstrated with a multi-scale modelling approach to map PFTs across the Peruvian Amazon. First they measured PFTs for all canopy trees in a network of field plots, capturing the physiological range of each trait. Next they trained regression models using each trait as the response variable, and the imaging spectroscopy data as features. These trait models were then applied to all airborne data, which were collected across gradients of elevation, geology and forest type. Finally, they modelled these traits continuously using the airborne-scale trait maps as the responses, and satellite measurements of ecosystem structure, ecosystem function, climate and topography as features. Since these traits vary widely within plots, and more so across the full study extent, the airborne-scale trait maps were a representative sample of local-scale trait variation across this high diversity region. The intermediate-scale maps provided more data to train the satellite-based models and, aggregated to the grain size of the satellite data, obviated problems of sampling effort and scale mismatch.

Applying these multi-scale modelling approaches could enable monitoring similar biodiversity patterns that have otherwise proven difficult to map over large extents. Though access to intermediate-scale data has been historically limited, it should increase with the launch of novel fine-grain sensors [Malenovský et al., 2012]. Monitoring intermediate-scale patterns could be further bolstered by expanding the scope of airborne mapping by groups like NEON's Airborne Operations Platform [Keller et al., 2008], DLR's Optical Airborne Remote Sensing platform (OpAiRS; [Baumgartner et al., 2012, Leutner et al., 2012], or the Carnegie Airborne Observatory (CAO; [Asner et al., 2012]). Linking field, airborne and spaceborne measurements could be used to map fine-scale patterns like species traits across large extents, generate intermediate-scale data to train and test satellite measurements, and link the distributions of community and ecosystem-scale patterns to species identities [Clark et al., 2005, Baldeck et al., 2015]. Furthermore, implementing large-scale airborne mapping efforts could be done at a fraction of the price of building and launching a satellite [Mascaro et al., 2014a].

2.6   Frontiers in Monitoring Biodiversity Change with EO

Global biodiversity monitoring systems hold great promise for advancing biodiversity science and conservation. These systems could be designed to forecast the rate, magnitude and geography of biodiversity change, identifying opportunities to mitigate human impacts on biological communities. EO can support monitoring systems by providing consistent and repeat assessments of biodiversity change; a unique global perspective of our changing biosphere. Applying concepts of pattern and scale in ecology to EO could link these fields in support of this vision. However, [Estes et al., 2018] found there is currently little overlap in the ecology literature between studies analysing field data and studies analysing EO data, highlighting the wide gap between these communities. And the problems presented by scaling dynamics (e.g. scale mismatches) have helped frame EO science as distinct from ecology, as subject to different rules and standards. Developing an ecologically based framework for monitoring biodiversity change with EO will require overcoming this distinction.

There are several key similarities between field and EO data: changing their grain size or extent fundamentally alters within and between-grain variation, there is rarely a single scale at which any pattern should be examined, and aggregating measurements to discrete domains of scale can constrain nonlinear responses to change. These similarities frame EO as an extension of field data; their differences are more in scale than they are in kind. Multi-scale analyses linking field and EO data support this, emphasising that targeted field collections are essential for mapping biodiversity patterns that are difficult to measure independently with EO. In the context of biodiversity monitoring, field data play three key roles: training EO to map novel biodiversity patterns; developing and testing forecasts of biodiversity change, and constraining the extents to which we can generalise patterns of change.

2.6.1   Ecologically Interpreting Sensor Data

One key challenge in measuring biodiversity patterns with EO is converting at-sensor measurements into biologically meaningful metrics of change (e.g. from at-sensor radiance to percent tree cover). This is often done empirically via calibration with field measurements. These calibrations require a lot of data; EO data dimensionality is often very high and the variation in biological communities that drives measurement variance is similarly high. However, it is difficult for any one research group to independently collect the field data necessary to capture this variation. One way to overcome this

challenge is to leverage open data.

Access to open biodiversity data has increased dramatically over the past decade [Kattge et al., 2011, Jetz et al., 2012, Metzger et al., 2013, Culina et al., 2018], as has access to open EO data [Nemani et al., 2011, Irons et al., 2012, Gorelick et al., 2017]. And though there are known spatial, temporal and taxonomic gaps in open biodiversity data [Beck et al., 2014, Geijzendorffer et al., 2016], extrapolating from incomplete measurements to fill these gaps is a key role for EO. Training global, multi-scale EO models using centralised and curated field data could provide baseline estimates of spatial biodiversity patterns that have been otherwise difficult to characterise. These baselines could be tested independently by researchers with improved local data and local knowledge, identifying opportunities to improve regional and global models. These analyses could spur modelling and data collection efforts to fill gaps, and to develop better forecasting tools. These are urgently needed in ecology [Dietze et al., 2018], and this theme is explored further in *Chapter 3*.

2.6.2   Reconciling Phenomenological and Process-based Models

Another key role for field data is to develop and test predictive, process-based models of temporal change. EO are uniquely suited for empirically monitoring change, especially for directly measurable patterns (e.g. disturbance) [Zhu et al., 2012, Cohen et al., 2016]. Yet forecasting change under conditions outside the range of historic variation (e.g. under novel climate and land-use scenarios) remains a challenge for EO. Furthermore, temporal lags between local environmental change and other scales of change (e.g. for species or community-scale patterns) can obfuscate efforts to identify the impacts of change [Essl et al., 2015]. Developing process-based models that couple temporal changes in EO to changes in other biodiversity patterns could address these issues [Korzukhin et al., 1996, Adams et al., 2013].

While there are many process-based EO models, and many process-based biodiversity models, we now have the technical capacity to link and test them using open data at multiple scales, identifying consensus models and key data gaps. Coupling process-based models with long-term, regularly updated and globally consistent measurements of change from EO could be used to develop early warning systems for identifying where species, communities and ecosystems will respond to change in novel ways, and may identify opportunities for science-driven mitigation (Daily 1999; Scholes et al . 2008). This theme is explored further in *Chapter 4*.

2.6.3 Bounding the Domains of Scale

Finally, field data are key for constraining how we generalise EO measurements and models of biodiversity change. One advantage of monitoring change with EO is that measurements are globally consistent; tree cover change can be mapped across tropical, temperate and boreal forests [Hansen et al., 2013, Sexton et al., 2013]. This enables other models of change that use tree cover data as features to be applied globally, such as the models of community composition in [Mendenhall et al., 2016]. This would be imprecise, however; the relationships between tree cover and community composition in tropical countrysides may not apply to timber plantations. In other words, this model is not stationary; the relationships between feature and response variables can change across the extent of the data [Hawkins, 2012]. The regions over which these relationships are stationary can be considered domains of scale, constraining the extents to which a model can generalise. It is currently difficult to identify these domains of scale with EO alone. Several algorithms can be employed to automate this task for EO (e.g. segmentation, clustering), but field data are key for interpreting and constraining these extents to biologically meaningful domains, and for testing their accuracy. Linking field and EO data to identify these domains of scale will be central to ecologically translating knowledge of local biodiversity patterns to regional and global scales.

2.6.4 Conclusion

After decades of work from biodiversity scientists, EO scientists and conservation groups, the stage is now set to establish ambitious, science-driven biodiversity monitoring systems. Consistent, repeat and globally available EO will play a key role in these systems. Scale is a central and unifying concept for biodiversity and EO sciences, and monitoring change with EO should be based on the principles and ecology of scale. Global biodiversity monitoring promises to expand our understanding of Earth's species, communities and ecosystems and, with luck, could help us discover the wisdom necessary to conserve them.

Chapter 3

Taxonomic Learning for Tree Species Mapping from Airborne Earth Observations

Christopher B. Anderson

3.1 Abstract

Biogeographers assess how species distributions and abundances affect the structure, function, and composition of ecosystems. Yet we face a major challenge: it is difficult to precisely map species locations across large landscapes. Novel Earth observations technologies could overcome this challenge for vegetation mapping. Airborne imaging spectrometers are able to measure plant functional traits at high resolution, which can be used to identify tree species in high resolution imagery. In this chapter I describe a trait- and taxonomy-based approach to species identification with imaging spectroscopy data, which was developed as part of an ecological data science competition. Using data from the National Ecological Observatory Network (NEON), I classified tree species using reflectance-based principal components rotation and decision tree-based machine learning models, approximating a morphological trait and dichotomous key classification inspired by botanical taxonomy.

The model received a rank-1 accuracy score of 0.919, and a cross-entropy cost score of 0.447 on the competition test data. Accuracy and specificity scores were high for all species, but precision and recall scores varied for rare species. PCA transformation improved accuracy scores compared to models trained using raw reflectance data, but outlier removal and data resampling exacerbated class imbalance problems. This taxonomy-informed approach accurately classified tree species using NEON data, reporting the best scores among data science competition participants. However, it failed to overcome several species mapping challenges like precisely classifying rare species. This method has been published under an open source, open access license, is designed for use with NEON data, and is publicly available to support future species mapping efforts.

3.2   Introduction

When you get down to it, biogeographers seek to answer two key questions: where are the species, and why are they there? Answering these simple questions continues to prove remarkably difficult. The former question belies a data gap; we do not have complete or unbiased information on where species occur, which is known as the "Wallacean shortfall" [Whittaker et al., 2005, Bini et al., 2006]. Addressing the latter question, however, does not necessarily require data; the drivers of species abundances and their spatial distributions can be derived from ecological theory and from niche use models [McGill, 2010]. But evaluating these theoretical predictions does require data. Testing generalized theories of species abundance distributions requires continuously-mapped presence and absence records for many individuals across large extents. And while field efforts can assess these spatial distribution patterns in fine detail, they are often restricted to small extents. Mapping organism-scale species distributions over large landscapes could help fill the data gaps that preclude addressing these key biogeographic questions. One remote sensing dataset holds the promise to do so for plants: airborne imaging spectroscopy.

Airborne imaging spectrometers measure variation in the biophysical properties of soils and vegetation at fine grain sizes across large extents [Goetz et al., 1985]. In vegetation mapping, imaging spectroscopy can measure plant structural traits, like leaf area index and leaf angle distribution [Broge and Leblanc, 2001, Asner and Martin, 2008], and plant functional traits, like growth and defense compound concentrations [Kokaly et al., 2009, Cavender-Bares et al., 2016]. These traits tend to be highly conserved within tree species, and highly variable between species (i.e., interspecific

trait variation is often much greater than intraspecific trait variation) [Wright et al., 2004, Townsend et al., 2007, Funk et al., 2017]. Trait conservation provides the conceptual and biophysical basis for species mapping with earth observations data. Airborne imaging spectroscopy has been used to map crown-scale species distributions across large extents in several contexts [Fassnacht et al., 2016]. These approaches have been applied in temperate [Baldeck et al., 2014] and tropical ecosystems [Hesketh and Sánchez-Azofeifa, 2012], using multiple classification methods [Feret and Asner, 2013], and using multiple sensors [Clark et al., 2005, Colgan et al., 2012, Baldeck et al., 2015]. However, this wide range of approaches has not yet identified a canonical best practice for tree species identification.

In the field, botanists can use plant morphological features and a dichotomous key to identify tree species. These features include reproductive traits (e.g., flowering bodies, seeds), vascular traits (e.g., types of woody and non-woody tissue), and foliar traits (e.g., waxy or serrated leaves). The dichotomous key hierarchically partitions species until each can be identified using a specific combination of traits, and the keys are aggregated to form a parsimonious taxonomy. Species classification with imaging spectroscopy is rather restricted in comparison; imaging spectrometers can only measure a subset of plant traits. This subset includes growth traits (e.g., leaf chlorophyll and nitrogen content), structural traits (e.g., leaf cellulose and water content), and defense traits (e.g., leaf phenolic concentrations and lignin content) [Lepine et al., 2016, Papeş et al., 2010, McManus et al., 2016]. Furthermore, the interspecific and intraspecific variation in this subset of traits is rarely known a priori; this precludes the use of a standard, accepted dichotomous key [Kichenin et al., 2013, Siefert et al., 2015].

Classifying species with imaging spectroscopy instead relies on distinguishing species-specific variation in canopy reflectance. Several confounding factors drive this variation [Ollinger, 2011, Lausch et al., 2016]:

1. Measurement conditions, like sun and sensor angles.

2. Canopy structure, including leaf area index and leaf angle distribution.

3. Leaf morphology and physiology (i.e., plant functional traits).

4. Sensor noise.

Measurement conditions and canopy structure tend to drive the majority of spectral variance; up to 79–89% of the signal is driven by within-crown spectral variation [Baldeck and Asner, 2014, Yao

et al., 2015]. Unfortunately, within-crown variation does not help distinguish between species. The spectral variation useful for discriminate between species is instead driven by differences in the plant functional traits expressed by a species [Asner et al., 2011, Martin et al., 2018]. Disentangling trait-based variation from measurement and structure-based variation is thus central to mapping species with imaging spectroscopy.

In this chapter, I describe an approach to tree species classification using airborne imaging spectroscopy data that builds on the above methods and takes a taxonomically-informed approach to data preprocessing and model selection in order to advance the discussion on best practices. This approach was developed as a submission to an Ecological Data Science Evaluation competition (ECODSE; `https://www.ecodse.org/`) sponsored by the National Institute of Standards and Technology. This competition provided participants with labeled tree crown data and airborne imaging spectroscopy data, collected by the National Ecological Observatory Network's Airborne Observation Platform (NEON AOP) [Kampe et al., 2010], which was used to identify tree crowns to the species level. The work described was submitted to the ECODSE competition under the team name of the Stanford Center for Conservation Biology (CCB), and has since been formalized as CCB-ID.

I first describe the CCB-ID approach to tree species classification using airborne imaging spectroscopy data, including both its successes and its shortcomings in the context of this competition, and close with key opportunities to improve future imaging spectroscopy-based species classification approaches. The goals of this work are to improve NEON's operational tree species mapping capacity, and to reduce technical and operational barriers to mapping plant biogeography over large extents.

3.3 Methods

CCB-ID classifies tree species using trait-based reflectance signal decomposition and decision tree-based machine learning models. This approach approximates a morphological trait and dichotomous key model to mapping species taxonomies [Godfray, 2007]. The first section describes data preprocessing procedures, including outlier removal and data decomposition procedures, which were used to reduce covariance and increase interpretability. The second section describes how the training data were resampled to reduce biases towards common species, as the training data contained a skewed distribution of species identities. The third section describes model selection, model training, and probability calibration, as an expansive series of metrics is required to evaluate classification models

trained on biased data. The fourth section describes the model performance metrics, and the final section describes two analyses performed post-ECODSE submission to evaluate the importance of signal decomposition and feature selection on model performance.

All data were provided by the ECODSE group (`https://www.ecodse.org/`) and are freely available from the NEON website (`https://neonscience.org`). All analyses were performed using the `Python` programming language (`https://python.org`) [Oliphant, 2007] and the following open source packages: `NumPy` (`http://numpy.org`) [der Walt et al., 2011], `scikit-learn` (`http://scikit-learn.org`) [Pedregosa et al., 2011], `pandas` (`https://pandas.pydata.org`) ([McKinney and Others, 2010], and `matplotlib` (`https://matplotlib.org`) [Hunter, 2007]. The python scripts used for these analyses were uploaded to a public GitHub repository (`https://github.com/earth-chris/ccb-id`), including a build script for a `Singularity` container to ensure the model is computationally reproducible [Kurtzer et al., 2017].

3.3.1   Data Preprocessing

The input NEON data were "Spectrometer orthorectified surface directional reflectance–flightline (NEON.DP1. 30008)" products. The canopy reflectance data were preprocessed using two steps: outlier removal and dimensionality reduction. In the outlier removal step, the reflectance data were spectrally subset, transformed using principal components analysis (PCA), then thresholded to isolate spurious values. First, reflectance values from the blue region of the spectrum $(0.38 - 0.49 \, \mu m)$ and from noisy bands $(1.35 - 1.43 \, \mu m$, $1.80 - 1.96 \, \mu m$, and $2.48 - 2.51 \, \mu m)$ were removed. These bands correspond to wavelengths dominated by atmospheric water vapor, and do not track variations in plant traits [Gao et al., 2009]. This reduced the data from 426 to 345 spectral bands.

Next, these spectrally-subset samples were transformed using PCA. The output components were whitened to zero mean and unit variance, and outliers were identified using a 3σ threshold. Samples with values outside of $\pm$ three standard deviations from the means (i.e., which did not fall within 99.7% of the variation for each component) for the first 20 principal components were excluded from analysis. These samples were expected to contain non-vegetation spectra (e.g., exposed soil), unusually bright or dark spectra, or anomalously noisy spectra [Féret and Asner, 2014]. The outlier-removed reflectance profiles for each species are shown in Fig. 3.1. Once the outliers were removed, the remaining spectral reflectance samples were transformed using PCA. This was not performed
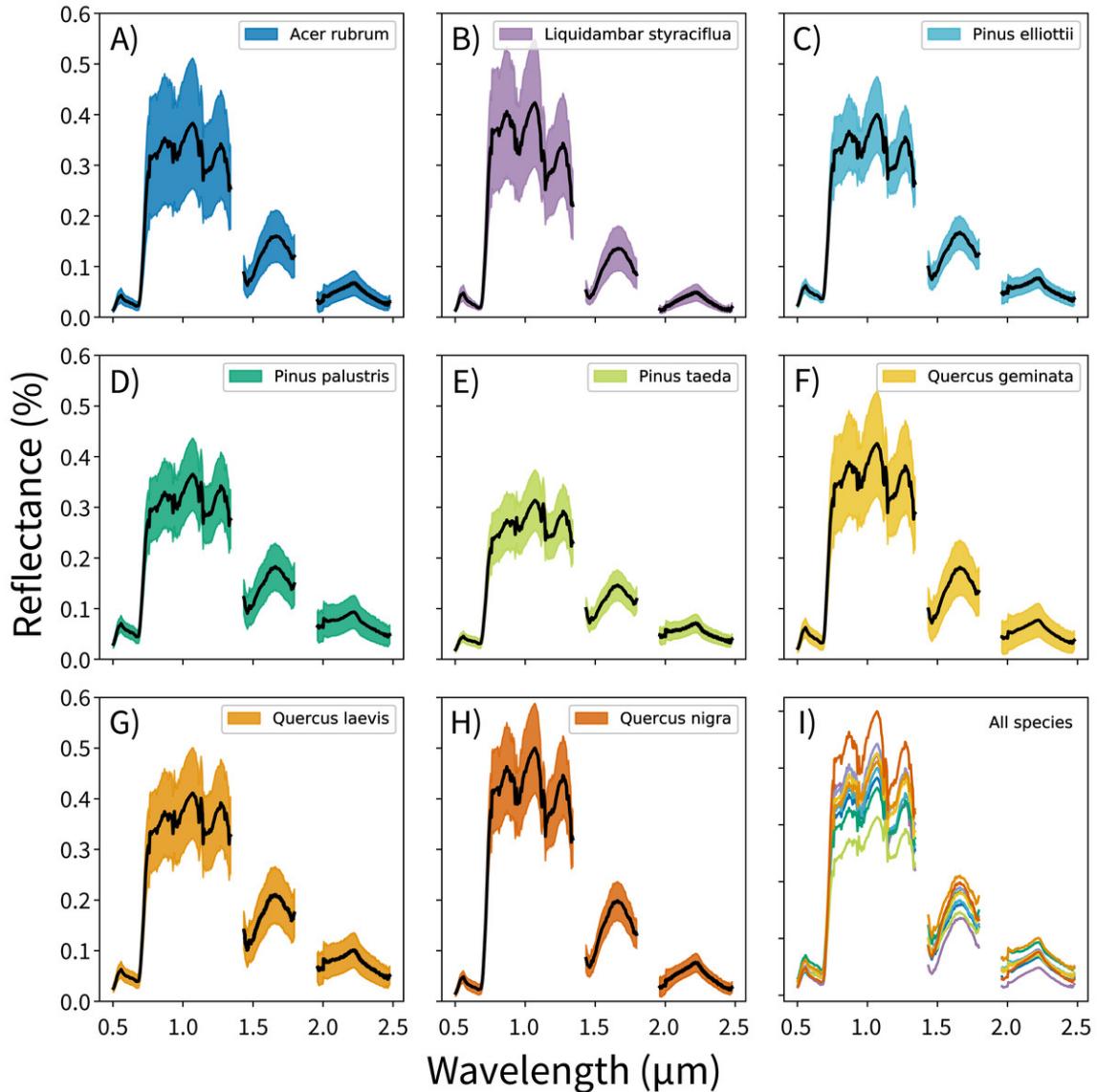
Figure 3.1: (A-H) Canopy reflectance profiles for eight tree species with mean reflectance values in black and $\pm$ 1 standard deviation in color. (I) Mean reflectance values for all species, with each color corresponding to the individual species panels.

on the already-transformed data from the outlier removal process, but on the outlier-removed, spectrally-subset reflectance data.

PCA transformations are often applied to airborne imaging spectrometer data to handle the high degree of correlation between bands, and these transformations are highly sensitive to input

feature variation [Jia and Richards, 1999]. Furthermore, transforming reflectance data into principal components can isolate the variation driven by measurement conditions from variation driven by functional traits. For example, brightness patterns alone can drive a high percentage of the variance in reflectance signals, which has nonlinear effects across the full spectrum, but PCA rotation isolates this variance into a single component. And though trait-based variation drives a small proportion of total reflectance signal, a single trait can be expressed in up to nine orthogonal components [Asner et al., 2015]. This is critical for distinguishing between species. After the transformation, the first 100 of 345 possible components were used as feature vectors for the classification models. This threshold was arbitrary; it was set to capture the majority of biologically-relevant components and to exclude the noisy components that explain a very low proportion of signal variance.

3.3.2 Class Imbalance

Class imbalance refers to datasets where the number of samples are not evenly distributed among classes. Imbalanced datasets are common in classification contexts, but can lead to problems if unaddressed. For example, training classification models with imbalanced data can select for models that overpredict common classes when the method of model selection doesn't weight misclassification penalties relative to the proportion that each class appears in the dataset. The ECODSE training data were imbalanced: after outlier removal, these data contained a total of 6,034 samples from nine classes (eight identified species, one "other species" class; Fig. 3.2). The most common species, *Pinus palustris*, contained 4,026 samples (66% of the samples) and the rarest species, *Liquidambar styraciflua*, contained 62 samples (1% of the samples).

These data were resampled prior to analysis to reduce the likelihood of overpredicting common species. Resampling was performed by setting a fixed number of samples per class, then undersampling or oversampling each class to that fixed number. This fixed number was set to 400 samples to split the difference of two orders of magnitude between the rarest and the most common classes. This number was arbitrary, but it was selected to approximate the number of per-species samples recommended in [Baldeck and Asner, 2014]. To create the final training data, classes with fewer than 400 samples were oversampled with replacement, and classes with more than 400 samples were undersampled without replacement. The final training data included 400 samples for each of the nine classes (3,600 samples total). Each sample contained a feature vector of the principal components
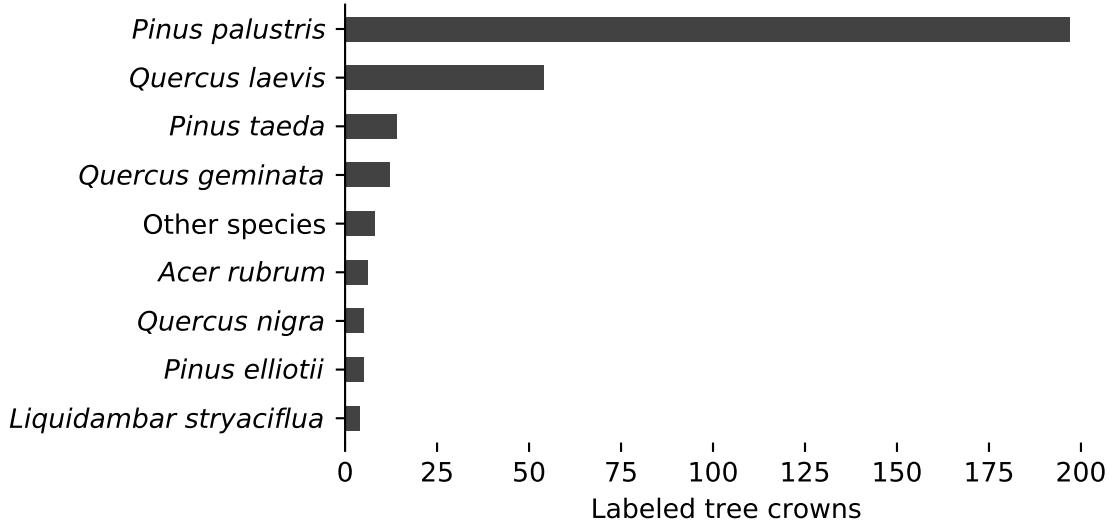
Figure 3.2: Number of training data crowns for each species. Class imbalance is common in ecology and biogeography, as species abundance distributions typically follow a log series distribution.

derived from the outlier removed, spectrally subset canopy reflectance data.

3.3.3   Model Selection and Probability Calibration

CCB-ID used two machine learning models: a gradient boosting classifier and a random forest classifier [Friedman, 2001, Breiman, 2001]. These models can fit complex, non-linear relationships between response and feature data, can automatically handle interactions between features, and have built-in mechanisms to reduce overfitting [Mascaro et al., 2014b]. They were selected because they perform well in species mapping contexts [Elith et al., 2008], in remote sensing contexts [Pal, 2005], and in conjunction with PCA transformations [Rodríguez et al., 2006]. Furthermore, these models are built as ensembles of decision trees, resembling the dichotomous key employed by botanists. Unlike a dichotomous key, these models are trained to learn where to split the data since the trait variation that distinguishes species was not known *a priori*.

These models were fit using hyper-parameter tuning and probability calibration procedures. Model hyper-parameters were tuned by selecting the parameters that maximized mean $F1$ scores in fivefold cross-validation using an exhaustive grid search. The F_1 score calculates the weighted average of model precision and recall (see *Model evaluation*), and maximizing F_1 scores during

model tuning reduces the likelihood of selecting hyper-parameters that overpredict common classes and underpredict rare classes. The following parameters were tuned for both models: number of estimators, maximum tree depth, minimum number of samples required to split a node, and minimum node impurity split threshold. The learning rate and node split quality criterion were also tuned for the gradient boosting and random forest classifiers, respectively. All samples were used for hyper-parameter tuning, and the best model hyper-parameters (i.e., the hyper-parameters that maximized mean F_1 scores in cross-validation) were used to fit the final models.

Prediction probabilities were calibrated (i.e., adjusted) after the final hyper-parameters were selected, as accurately characterizing prediction probabilities is essential for error propagation and for assessing model reliability. Well-calibrated probabilities should scale linearly with the true rate of misclassification (i.e., model predictions should not be under or overconfident). Some decision tree ensemble methods, such as random forest, tend to be poorly calibrated, however. Since this type of ensemble averages predictions from a set of weak learners—which individually have high misclassification rates but gain predictive power as an ensemble—model variance can skew high probabilities away from one, and low probabilities away from zero. This results in sigmoid-shaped reliability diagrams [DeGroot and Fienberg, 1983, Niculescu-Mizil and Caruana, 2005].

To reduce these biases, prediction probabilities were calibrated using sigmoid regression for both the gradient boosting and random forest classifiers. The data were first randomly split into three subsets: model training (50%, or 200 samples per class), probability calibration training (25%, or 100 samples per class), and probability calibration testing (the remaining 25%). Each classifier was fit using the model training subset and the tuned hyper-parameters. Prediction probabilities were calibrated with sigmoid regression using the probability training subset, and internal threefold cross-validation to assess the calibration. Calibrated model performance was assessed using the holdout test data. After these assessments, the final models were fit using the model training data, then calibrated using the full probability training and testing data (i.e., the full 50% of samples not used in initial model training).

3.3.4 Model Evaluation

During model training, performance was assessed on a per-sample basis using model accuracy and log loss scores. Model accuracy calculates the proportion of correctly classified samples in the test data
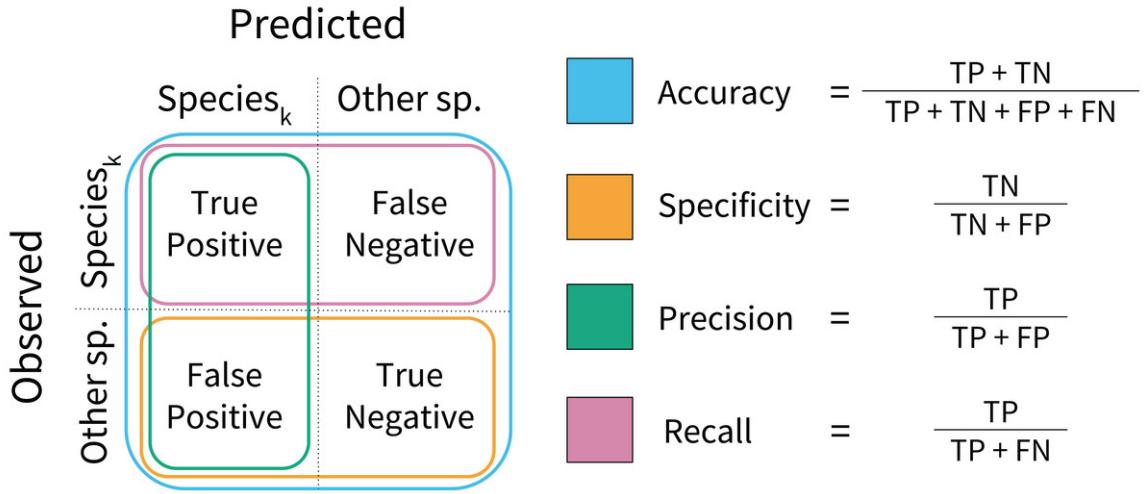
Figure 3.3: Summary of the classification model metrics calculated on per-pixel and per-species bases. A confusion matrix was computed for each species, and each metric was calculated in a one-vs-all fashion.

(Fig. 3.3), and high model accuracy scores are desirable. Log loss assesses whether the prediction probabilities were well calibrated, penalizing incorrect and uncertain predictions. Low log loss scores indicate that misclassifications occur at rates close to the rates of predicted probabilities. During model testing, performance was assessed using rank-1 accuracy and cross-entropy cost [Marconi et al., 2018]. Rank-1 accuracy was calculated based on which species ID was predicted with the highest probability. The cross-entropy score is similar to the log loss function, but was scaled using an indicator function. These can be interpreted in similar ways to accuracy and log loss scores; high rank-1 accuracy and low cross-entropy scores are desirable [Hastie et al., 2009].

Secondary model testing metrics were calculated for each species using the test data. These included model specificity, precision, and recall (Fig. 3.3). These metrics reveal model behavior that accuracy scores may obscure. Specificity assesses model performance on non-target species, penalizing overprediction of the target species (i.e., a high number of false positives). Precision also penalizes overprediction, but assesses the rate of overprediction relative to the rate of true positive predictions. Recall calculates the proportion of true positive predictions to the total number of positive observations per species. Higher values are desirable for each. These metrics were calculated to aid interpretation, but were not used to formally rank model performance.

Performance during model training was assessed at the sample scale, meaning the model performance metrics were calculated on every pixel (i.e., sample) in the training data. However, the competition evaluation metrics were calculated using crown-scale prediction probabilities, so performance metrics were calculated after aggregating each pixel from individual trees to unique crown identities. To address this scale mismatch, prediction probabilities were first calculated for each sample in a crown using both gradient boosting and random forest models, then the sample-scale probabilities were then averaged by crown.

3.3.5 Decomposition Analysis

Two analyses were performed to assess how PCA transformations affected model performance. Prior to these analyses, I bootstrapped the original model fits to assess model variance. I then compared these bootstrapped fits to models trained with the spectrally-subset reflectance data instead of the PCA transformed data. This was done to evaluate the change in model performance attributed to the PCA rotation. Next, I compared models trained using a varying number of principal components. These models were trained using $n_{pcs} \in (10, 20, \ldots, 345)$ as the input features, with 345 being the maximum number of potential components after spectral subsetting. These comparisons assessed whether the PCA transformations improved model performance and how changing the amount of spectral variation in the feature data affected performance. These were each bootstrapped 50 times to derive an unbiased estimate of the variance in each model.

3.4 Results

CCB-ID performed well according to the ECODSE competition metrics, receiving a rank-1 accuracy score of 0.919, and a cross-entropy cost score of 0.447 on the test data. These were the highest rank-1 accuracy and the lowest cross-entropy cost scores among participants. Other methods reported rank-1 accuracy scores from 0.688 to 0.88 and cross-entropy scores from 0.877 to 1.448 [Marconi et al., 2018]. A confusion matrix with the classification results is reported in Fig. B.2. In addition to the high rank-1 accuracy and low cross-entropy cost scores, the CCB-ID model performed well according to the secondary crown-scale performance metrics. These secondary metrics calculated a mean accuracy score of 0.979, mean specificity score of 0.985, mean precision score of 0.614, and mean recall score of 0.713 across all species. The per-species secondary metrics are summarized in Fig. 3.4. These
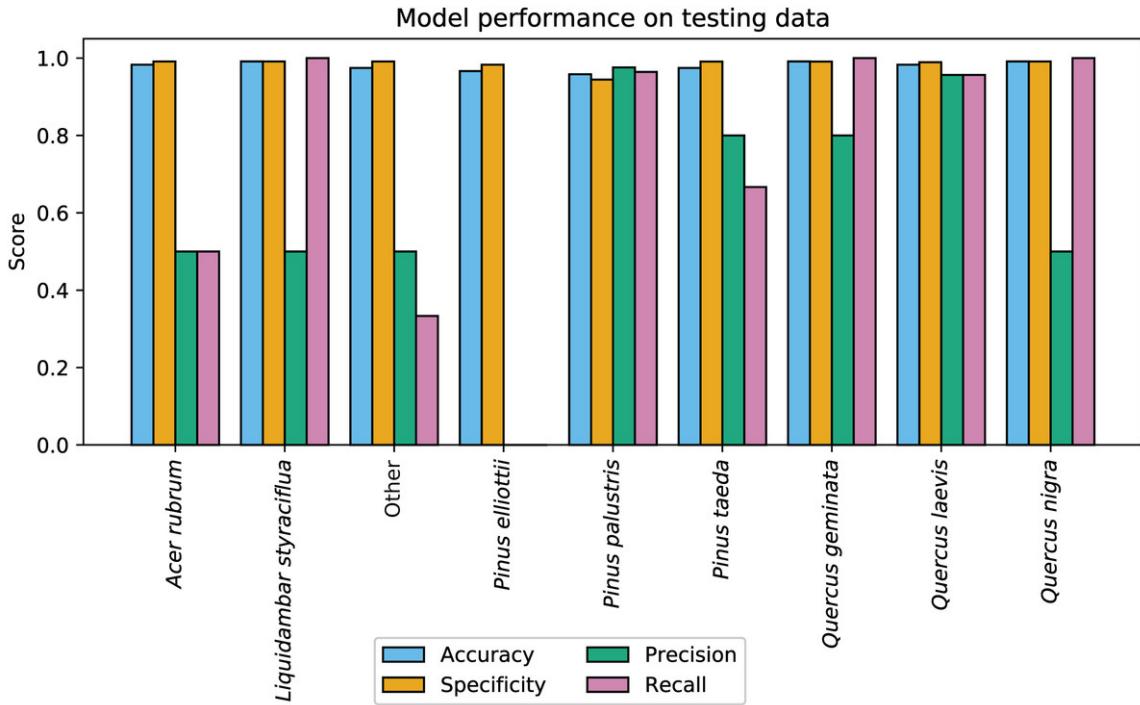
Figure 3.4: Per-species secondary model performance metrics applied to test data, calculated using the binary confusion matrix reported in Fig. B.2. Metrics weighted by the true negative rate (i.e., accuracy and specificity) were high for all species since the models correctly predicted the most common species, *Pinus palustris*. However, metrics weighted by the true positive rate (i.e., precision and recall) were much more variable since there were fewer than six observed crowns for seven of the nine species (*P. palustris* and *Quercus laevis* had 84 and 23 crowns, respectively). This penalized misclassifications of rare species. These metrics were also recalculated using the continuous per-crown prediction probabilities, which can be found in Fig. B.1

results were calculated using the categorical classification predictions (i.e., after assigning ones to the species with the highest predicted probability, and zeros to all other species). Probability-based classification metrics are reported in Fig. B.1.

During model training, outlier removal excluded 797 samples from analysis. A total of 264 of the 797 samples (33%) removed from analysis were from *P. palustris*, while the remaining 533 samples (67%) were from non-*P. palustris* species. Outlier removal disproportionately excluded samples from uncommon species; 45% of samples from *L. styraciflua*, the rarest species, were removed. After outlier removal, the first principal component contained 78% of the explained variance. However, this component did not drive model performance; it ranked 7th and 11th in terms of ranked feature importance scores for the gradient boosting and random forest classifiers. Model accuracy scores,
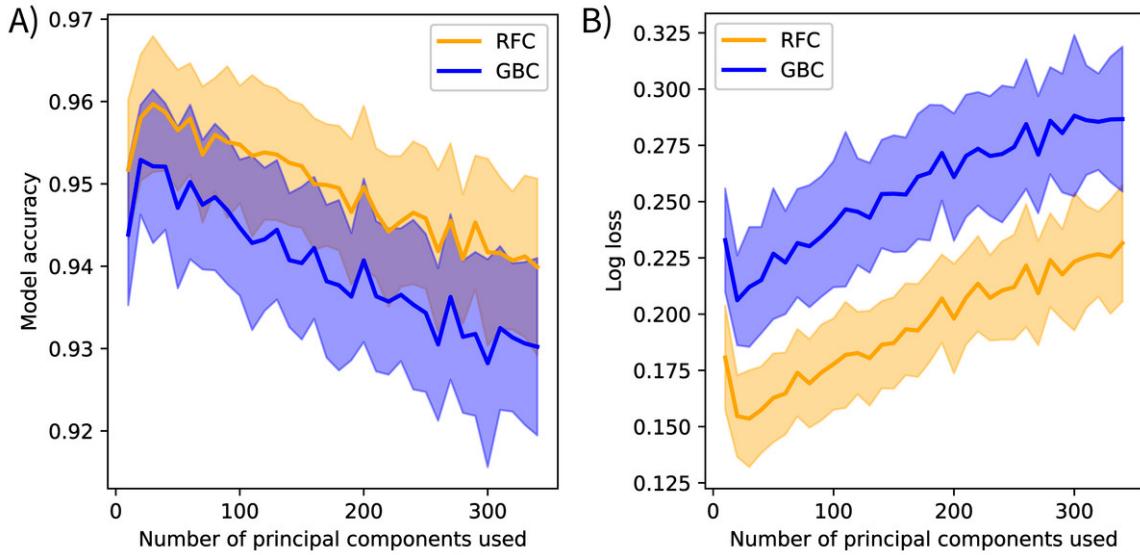
Figure 3.5: Mean (solid) and standard deviation (shaded) of (A) model accuracy and (B) log loss scores as a function of increasing spectral variance for each classification method. Scores were calculated on holdout data from the training set, not the competition test data. Using all spectral variance (i.e., all principal components) as model features decreased model performance. Both random forest classifiers (RFC) and gradient boosting classifiers (GBC) were used.

calculated on a sample basis (i.e., not by crown) using the 25% training data holdout, were 0.933 for gradient boosting and 0.956 for random forest. Log loss scores, calculated prior to probability calibration, were 0.19 for gradient boosting, and 0.47 for random forest. After probability calibration, log loss scores were 0.24 for gradient boosting and 0.16 for random forest. The per-class secondary metrics reported a mean specificity score of 0.987, mean precision score of 0.908, and mean recall score of 0.907 across all species.

The post-submission analyses found PCA transformations improved model accuracy. Models fit using the original methods calculated mean bootstrapped accuracy scores of 0.944 ($s.d. \pm 0.009$) for gradient boosting and 0.955 ($s.d. \pm 0.008$) for random forest. Models fit using the spectrally-subset reflectance data as features calculated mean accuracy scores of 0.883 ($s.d. \pm 0.012$) for gradient boosting and 0.877 ($s.d \pm 0.011$) for random forest, and mean log loss scores of 0.46 ($s.d. \pm 0.03$) for gradient boosting and 0.48 ($s.d. \pm 0.03$) for random forest. Mean model accuracies declined and mean log loss scores increased after including more than 20 components as features for the models fit using varying numbers of principal components (Fig. 3.5).

3.5 Discussion

CCB-ID accurately classified tree species using NEON imaging spectroscopy data, reporting the highest rank-1 accuracy score and lowest cross-entropy cost score among ECOSDE participants. These scores compare favorably to other imaging spectroscopy-based species classification efforts, as reviewed by [Fassnacht et al., 2016]. These crown-scale test results highlight the technical and conceptual potential of species mapping methods that approximate botanical and taxonomic approaches to classification. However, this method failed to overcome several well-known species mapping challenges, like precisely identifying some rare species. Below I discuss some key takeaways and suggest opportunities to improve future imaging spectroscopy-based species classification approaches.

3.5.1 Class Imbalance in Ecological Contexts

The high per-species accuracy scores indicate a high proportion of correctly classified crowns in the test data. However, accuracy can be a misleading metric in imbalanced contexts. Since seven of the nine classes had six or fewer crowns in the test data (out of 126 total test crowns), classification metrics weighted by the true negative rate (i.e., accuracy and specificity) were expected to be high if the majority class were correctly predicted. Metrics weighted instead by the true positive rate (i.e., precision and recall) showed much higher variation across rare species, as a single misclassification greatly alters these metrics when there are few observed crowns (Fig. 3.4). Due to the small sample size, it is difficult to assess if these patterns portend problems at larger scales. For example, there were two observed *Acer rubrum* crowns in the test data, yet only one was correctly predicted. Was the misclassified crown an anomaly? Or will this low precision persist across the landscape, predicting *A. rubrum* occurrences at half its actual frequency? The latter seems unlikely, in this case; the low cross-entropy and log loss scores suggest misclassified crowns were appropriately uncertain when assigning the wrong label (B.1). However, since airborne species mapping is employed to address large-scale ecological patterns where model precision is key (e.g., in biogeography, macroecology, and biogeochemistry), we should be assessing classification performance on more than one or two crowns per species.

Comparing model performance between and within taxonomic groups revealed notable patterns. *Quercus* and *Pinus* individuals (i.e., Oaks and Pines) accounted for 120 of the 126 test crowns and there was high fidelity between them. Only one Quercus crown was misclassified as *Pinus*, and two

Pinus crowns were misclassified as *Quercus*. From a botanical perspective, this makes sense; these genera exhibit very different growth forms (i.e., different canopy structures and foliar traits), and should thus be easy to distinguish in reflectance data. However, within-genus model performance varied between *Quercus* and *Pinus*. *Quercus* crowns were never misclassified as other *Quercus* species, yet there were several within-*Pinus* misclassifications.

This may be because *Quercus* species tightly conserve their canopy structures and foliar traits [Cavender-Bares et al., 2016], while *Pinus* species may express trait plasticity. *Pinus* species maintain similar growth forms (i.e., their needles grow in whorls bunched through the canopy), perhaps limiting opportunities to distinguish species-specific structural variation. Furthermore, they are distributed across the varying climates of the southern, eastern, and central United States, suggesting some degree of niche plasticity. If this plasticity is expressed in each species' functional traits, then convergence among species may then preclude trait-based classification efforts. Quantifying the extent to which foliar traits are conserved within and between species and genera will be essential for assessing the potential for imaging spectroscopy to map community composition across large extents [Violle et al., 2012, Siefert et al., 2015].

3.5.2   Trait-based Interpretations of Imaging Spectroscopy Data

The post-submission analyses revealed several notable patterns. First, PCA transformation significantly increased mean model accuracy scores by 7-8% compared to the spectrally-subset reflectance data. I suspect this is because the models could focus on the spectral variation driven by biologically meaningful components instead of searching for that signal in the continuous reflectance spectrum, where the majority of variation is driven by abiotic factors. The low feature importance scores of the first principal component support this interpretation. The first component in reflectance data is typically driven by brightness —not an indicator of interspecific variation—and contained 78% of the explained reflectance variance, but ranked low in feature importance for both models. This preprocessing transformation approximates the "rotation forest" approach developed by [Rodríguez et al., 2006], who found PCA preprocessing improved tree-based ensemble models in virtually every context it was applied. They suggested retaining all components to maintain the full dimensionality of the input data. However, the component-based sensitivity analysis showed model accuracy decreased when including more than the first 20 components (Fig. 3.5). Since there are many components that

include non-biological information, like brightness or sensor noise, there results suggest that using all components overfits to spurious signals in this feature-rich dataset. Performing feature selection on transformed data will likely help to overcome this.

Feature selection has been applied to reflectance data to find the spectral features that track functional trait variation [Feilhauer et al., 2015], which should help identify the trait-based principal components that discriminate between species. Furthermore, other transformation methods may be more appropriate than PCA; principal components serve only as proxies for functional traits in this context. I expect transforming reflectance data directly into trait features, further extending the analogy of a taxonomic approach to classification, would improve species mapping accuracy, improve model interpretability, and further define the mechanistic and biophysical basis for species mapping with imaging spectroscopy.

3.5.3 Overcompensating for Rarity

Despite the successes of CCB-ID, there were several missteps in model design and implementation. For example, outlier removal and resampling were employed to reduce class imbalance problems but may instead have exacerbated them. First, the PCA-based outlier removal excluded samples based on deviation from the mean of each component. However, since the transformations were calculated using imbalanced data, the majority of the variance was driven by variation in the most common species. This means outlier removal excluded samples that deviated too far from the mean-centered variance weighted by *P. palustris*. Indeed, 533 of the 797 samples excluded from analysis (67%) were from non-*P. palustris* species (which comprised only 37% of the full dataset). This removed up to 45% of samples from the rarest species (*L. styraciflua*), reducing the spectral variance these models should be trained to identify. This suggests outlier removal should either be skipped or implemented using other methods (e.g., using spectral mixture analysis to identify samples with high soil or non-photosynthetic vegetation fractions) to reduce imbalance for rare species.

Data resampling further exacerbated class imbalance. By setting the resampling threshold an order of magnitude above the least sampled class, the rarest species were oversampled nearly 10-fold in model training. This oversampling inflated per-class model performance metrics by double-counting (or more) correctly classified samples for oversampled species. These metrics were further inflated as a result of how the train/test data were split. The split was performed after resampling, meaning the

train/test data for oversampled species were likely not independent. This invalidated their use as true test data, overestimating performance during model training. This was bad practice. Undersampling the common species was also detrimental. Excluding samples from common species meant the models were exposed to less intraspecific spectral variation during training. This is a key source of variation the models should recognize. Excluding this spectral variation made it more difficult for the models to distinguish inter and intraspecific variation. Assigning sample weights (e.g., proportional to the number of samples per class) and using actually independent holdout data could overcome these issues.

3.6   Conclusion

Airborne imaging spectrometers can map tree species at crown scales across large areas, and these data are now publicly available through NEON's open data platform. However, there is currently no canonical imaging spectroscopy-based species mapping approach, limiting opportunities to explore key patterns in biogeography. This taxonomic learning approach was developed to address this gap and to further the conversation on best practices for species mapping. CCB-ID performed well within the scope of the ECODSE competition, reporting the highest rank-1 accuracy and lowest cross-entropy scores among participants. Yet further testing is necessary to identify whether this method can scale across multiple sites or to other regions, including high diversity forests. I hope CCB-ID will be used by others to improve future species mapping efforts in pursuit of answers to biogeography's great mysteries of where the species are, and why they are there.

Chapter 4

Niche Use and Niche Conservation Under Climate, Habitat and Resource Constraints for Two Global Arbovirus Vectors

Christopher B. Anderson[1,2], Erin A. Mordecai[1], Meghan E. Howard[1], Luis E. Fernandez[3,4], Ricardo Gamboa[5], Marcelo Guevara[6], Morgan P. Kain[1,6], Andres G. Lescano[7], Lisa Mandle[6], Stephanie Montero[7], Mileyka Santos[8], Jeffrey R. Smith[1,2], Adrian Vogl[6], and Gretchen C. Daily[1,2,6,7]

[1]Department of Biology, Stanford University, Stanford, USA

[2]Center for Conservation Biology, Stanford University, Stanford, USA

[3]Centro de Innovación Científica Amazónica (CINCIA), Puerto Maldonado, Perú

[4]Center for Energy, Environment, and Sustainability, Wake Forest University, Winston-Salem USA

[5]School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Perú

[6]Natural Capital Project, Stanford University, Stanford, USA

[7]Woods Institute for the Environment, Stanford University, Stanford, USA

[8]Department of Medical Entomology, Gorgas Memorial Institute of Health Studies, Panama City, Panamá

4.1 Abstract

Mosquitoes are expected to shift their geographic distributions with rising temperatures, urbanization, agricultural expansion and human population growth. As ectotherms, temperature responses are mechanistically understood for many mosquito arbovirus vectors. But the responses to other environmental changes are not. How do these covarying patterns of change interact to determine the biogeography of these vectors? Here, we quantified how three distinct dimensions of the vector niche—climate, habitat, and resource constraints—interact to determine the distributions of two mosquito species, *Aedes aegypti* and *Ae. albopictus*, which transmit dengue, chikungunya, Zika, and viruses across Latin America and the Caribbean. When considered independently from other drivers, resource constraints (i.e., access to blood meals) best predicted the realized niche for *Ae. aegypti*, while temperature constraints best predicted niche patterns for *Ae. albopictus*. Both vectors occurred disproportionately in areas with high mean and low variance in temperature throughout the year, revealing strong niche preferences for high temperatures and thermal stability. *Ae. aegypti* was more frequently observed at warmer temperatures (mean = $31.0°C$) than *Ae. albopictus* (mean = $29.1°C$), consistent with mechanistic predictions of vector-specific thermal optima. *Ae. aegypti* occurred in areas of higher population density (mean = $632.2\,people\,km^{-2}$) than *Ae. albopictus* (mean = $363.6\,people\,km^{-2}$), which tended to occur in areas of higher livestock density (mean = $4.4\,animals\,km^{-2}$). Resource use patterns were consistent across Mesoamerica, South America, and the Caribbean for both vectors, while climate and habitat patterns were region-specific, suggesting *Aedes* distribution patterns may be constrained by access to blood meals.

4.2 Introduction

The global burden of dengue is increasing worldwide, with the number of symptomatic infections doubling every decade since 1990 [Stanaway et al., 2016]. The geography of transmission is expected to shift under climate change [Bhatt et al., 2013, Campbell et al., 2015], increasing in temperate areas and decreasing in areas that will become too hot to support sustained viral transmission [Mordecai et al., 2019, Ryan et al., 2019]. These shifts may be especially pronounced in Latin America and the Caribbean [Shepard et al., 2011, Tapia-Conyer et al., 2012], which have seen rising temperatures, shifting precipitation regimes, and rapid forest loss, each of which promote dengue transmission [Locatelli et al., 2011, Marengo et al., 2011, Collins et al., 2013, Nobre et al., 2016]. Reduced vector control practices after the 1960s [Gubler, 2002, Gómez-Dantés and Willoquet, 2009] and human-assisted migration of mosquitoes also drove reintroductions and new invasions across the region [Knudsen, 1995, Rossi et al., 1999, Tatem et al., 2006, Ortega-Morales and Rodríguez, 2016]. Recent events indicate short-term increases in transmission as well: Perú announced a

dengue emergency for three Amazonian departments in February 2020 [Ministerio de Salud, Perú, 2020, U.S. Embassy Lima, 2020] and Brazil reported nearly 1.5 million dengue cases in the 2019 dry season alone—more than all cases reported in 2017 and 2018 combined [Brazil, 2019, Pan American Health Organization / World Health Organization, 2020].

The globally invasive mosquito species *Aedes aegypti* and *Ae. albopictus* are the primary vectors of dengue, as well as Zika and chikungunya viruses, which also caused explosive epidemics in Latin America and the Caribbean in 2016-2017 and 2014-2015, respectively. With no vaccines widely available for these viruses (collectively called arthropod-borne viruses or "arboviruses"), forecasting and intervention are key mitigation strategies [Tapia-Conyer et al., 2012, Altizer et al., 2013, Castro et al., 2019]. As transmission is driven by species interactions between humans and mosquitoes, and moderated by environmental conditions [Hopp and Foley, 2001, Paaijmans et al., 2013, Mordecai et al., 2019], transmission forecasts depend to a large degree on how well we can characterize human-vector-environment interactions.

There are two main approaches to modeling how humans and the environment determine mosquito biogeography [Tjaden et al., 2018]. Mechanistic models explicitly define the physiological and ecological processes driving vector life cycles and infection patterns [Costa et al., 2010, Carrington et al., 2013, Mordecai et al., 2013]. Quantifying the temperature dependence of these processes has been a central focus for ecological epidemiologists, with both lab experiments and mathematical models predicting that mosquito life history and interaction traits vary directly and nonlinearly with temperature [Brown et al., 2004, Mordecai et al., 2017, Shocket et al., 2018]. However, it is unclear whether these thermal responses, alone or in combination with other drivers, predict vector biogeography with the same precision [Tjaden et al., 2018]. Alternatively, species distribution models (SDMs) characterize biogeographic patterns of where species are present, using occurrence records and environmental data to quantify a species' realized niche (the conditions of occupied habitat) and to predict its fundamental niche (potentially suitable habitat), though these methods eschew mechanism for predictive power [Grinnell, 1917, Soberon and Townsend Peterson, 2005, Wiens et al., 2009]. Global SDMs for *Ae. aegypti* and *Ae. albopictus*, and for the arboviruses they transmit, report that temperature and precipitation are the predominant drivers of vector biogeography and are the best predictors of epidemic potential [Brady et al., 2014, Liu-Helmersson et al., 2014, Kraemer et al., 2015a], corroborating mechanistic models. Recent studies have merged these approaches by applying lab-derived thermal responses to climate-driven niche maps to forecast spatial transmission patterns [Tjaden et al., 2017, Mweya et al., 2016].

Despite this consistency between mechanistic and statistical methods, it's not yet clear that climate alone sufficiently characterizes the mosquito vector niche. Spatially-explicit cross-validation studies suggest it may not [Liu et al., 2020]. Multiple climate-driven SDMs for *Ae. aegypti* and *Ae. albopictus*, trained on data from

Asia, Europe and North America, have failed to generalize to Latin America where both vectors are invasive, suggesting that vector-climate relationships are not consistent across continents [Medley, 2010, Carlson et al., 2016, Pech-May et al., 2016]. These results have been interpreted as evidence of niche evolution; that each vector is adapting to new climates as it invades. But niche evolution is a complex process, and climate constraints only characterize one dimension of the vector niche [Wiens et al., 2010, Hortal et al., 2015, Diniz-Filho and Bini, 2019]. Moreover, no direct evidence exists to support the idea that vector thermal constraints are evolving. Is it instead possible that other conserved ecological mechanisms, like habitat or resource constraints, better explain spatial invasion patterns?

Aedes habitat constraints are typically characterized by land cover, with *Ae. aegypti* often found in vegetated urban areas [Troyo et al., 2009, Jansen and Beebe, 2010] and *Ae. albopictus* found in rural and agricultural areas [Knudsen, 1995, Tsuda et al., 2006], though the urban/rural dichotomy does not always hold [Li et al., 2014]. In global models, the importance of climate overshadows land cover when solely characterized by simple remote sensing metrics [Kraemer et al., 2015a, Carlson et al., 2016]. Yet land cover better describes vector distributions when represented by descriptive metrics like vegetation phenology, tree cover, or building density, which mechanistically map to access to sugar feeding resources, natural breeding sites, or urban breeding sites, respectively [Martinez-Ibarra et al., 1997, Peterson et al., 2005, Vanwambeke et al., 2007, Troyo et al., 2009, Landau and van Leeuwen, 2012, Li et al., 2014]. These hematophagous vectors are also reproductively constrained by access to blood meals [Padmanabha et al., 2012]. Humans are the primary blood meal source for both vectors, but *Ae. albopictus* is less specific and will often feed on domesticated mammals and on wildlife [Gratz, 2004, Delatte et al., 2008]. Despite the outsized importance of resource constraints in other ecological contexts, they are rarely quantified when mapping *Aedes* niche use.

We hypothesize that spatial invasion patterns that are unexplained by climatic constraints may instead be explained by habitat or resource constraints. We tested this by quantifying niche preferences and evaluating evidence for niche conservation in *Ae. aegypti* and *Ae. albopictus* across Latin America and the Caribbean, where both vectors are invasive. We asked:

1. To what degree do climate, habitat, and resource constraints predict spatial patterns of niche use for *Ae. aegypti* and *Ae. albopictus* at continental scales?

2. How do statistically-derived niche preferences compare to direct lab and field observations?

3. Are climate, habitat, and resource preferences conserved across Latin America and the Caribbean for each vector?

We trained presence-only SDMs using 11 mechanistically-based covariates to quantify the importance of these niche dimensions in characterizing each vector's realized niche. We compared our results to lab-derived

thermal responses and to two mosquito abundance field surveys to corroborate experimental and *in situ* patterns. To evaluate evidence of niche evolution, we tested whether models trained in one region predicted spatial distributions in other regions (i.e., whether niche preferences are conserved). We evaluated model performance using area under the receiver operator curve (AUC), a metric of separability, which calculates the probability that model predictions distinguish between suitable habitat and the background (i.e., a bias-adjusted random geographic sample). We adjusted for sampling bias by assuming sampling effort was biased towards urban areas by applying the target group background selection method, using a global nightlights dataset and non-Aedes occurrence records from the Culicidae family [Phillips et al., 2009, Merow et al., 2013]. This is distinct from studies that solely restrict background sample selection to vector-specific envelopes of temperature suitability [Brady et al., 2014, Kraemer et al., 2015a]. We exclusively used satellite-derived covariates, which provide spatially continuous and regularly-spaced measurements instead of interpolated weather station data, which are sparse across Latin America and the Caribbean [Fick and Hijmans, 2017].

4.3   Methods

Species distribution modeling is based on the Grinellean niche concept: the environmental conditions that allow individuals of a species to survive and reproduce will constrain the distributions of those species [Grinnell, 1917, Wiens et al., 2009]. The inputs to these models are spatially explicit species occurrence records and gridded environmental covariates, which we gathered and derived from publicly-accessible datasets (Fig. 4.1). We also included a bias assessment in our models, described in-depth below, as *Aedes* records are often collected in populated areas in efforts to map human disease risk. We sampled background covariates in proportion to the sampling effort of the occurrence data (i.e., we assessed suitability relative to the environmental conditions where sampling was most likely to occur). All spatial data were projected to a Molleweide global equal area, $1\,km^2$ grid prior to analysis (EPSG:54009), all map figures made in web mercator (EPSG:3857), all scripts can be found on GitHub (https://github.com/earth-chris/aedes-americas), all species distribution models were trained using the Maxent software [Steven J. Phillips, Miroslav Dudík, Robert E. Schapire, 2017], and all spatial analysis was performed in Python or on Google Earth Engine [Gorelick et al., 2017].

4.3.1   Occurrence Records and Environmental Covariate Data

We analyzed 6,317 *Aedes aegypti* occurrence records and 3,629 *Ae. albopictus* records from two global datasets: the Global Biodiversity Information Facility (GBIF; https://gbif.org) and from a global synthesis by [Kraemer et al., 2015b]. All data were filtered to contain only records in Latin America and the Caribbean
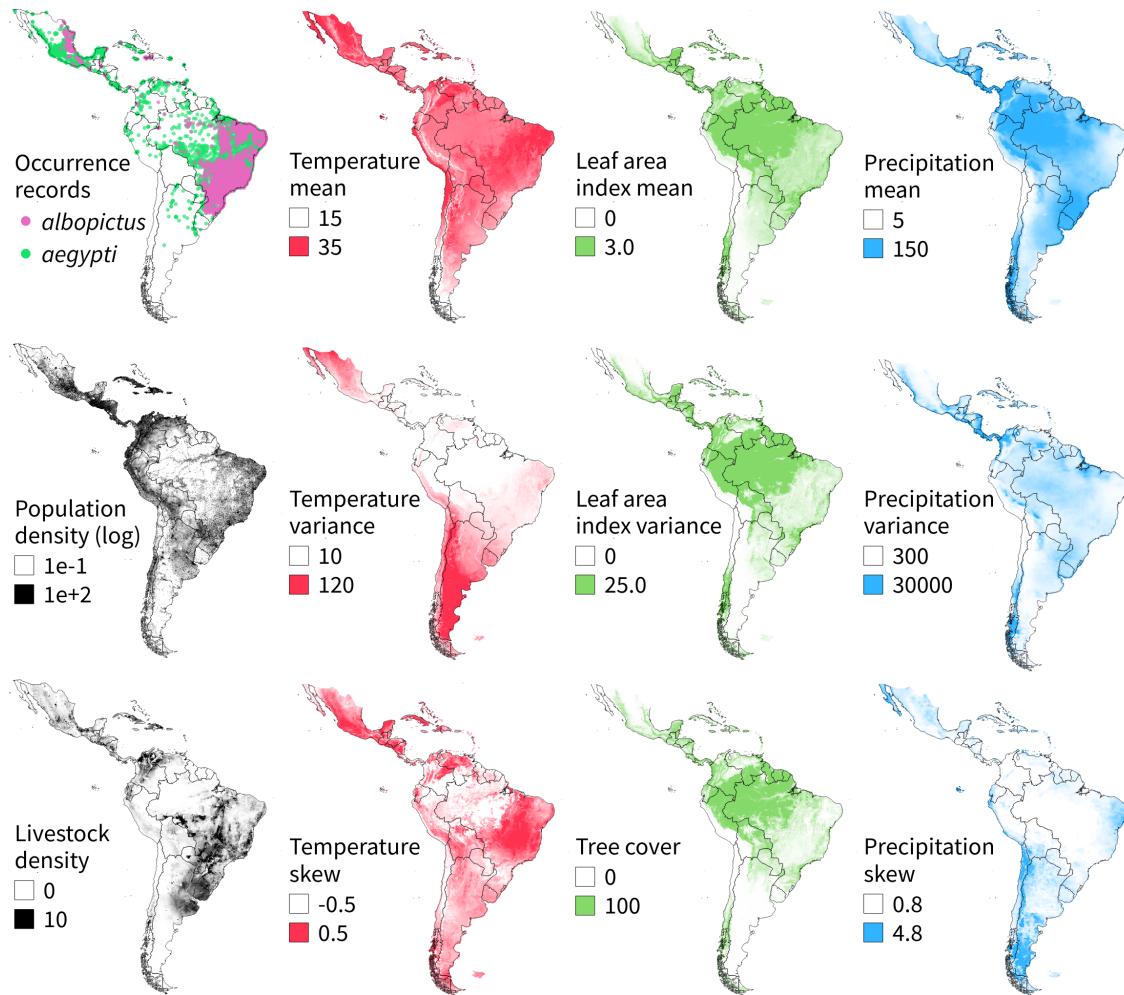
Figure 4.1: Species occurrence records and environmental covariates show continental-scale niche use patterns that determine *Aedes* distributions. The human population density map (black, top) was $log_1 0$ transformed from units of *people ha^{-1}*. The livestock density map (black, bottom) is in units of *animals km^{-2}*. Temperature patterns (red) are reported in $^{\circ}C$. Leaf area index patterns (green, top) are reported in units of $m^2 m^{-2}$, while tree cover (green, bottom) is reported in %. Precipitation patterns (blue) are reported in units of $mm mo^{-1}$.

from 2000 to 2017. Raw *Ae. aegypti* data from GBIF contained 5,648 records [GBIF, 2018a], and the raw *Ae. albopictus* data contained 3,452 records [GBIF, 2018b]. All GBIF data were cleaned to include only published records, to exclude points with > 10 km reported spatial uncertainty, and to include only points recorded by human observation. The raw *Ae. aegypti* data from Kraemer et al. contained 1,067 records, and the raw *Ae. albopictus* data contained 159 records, which were quality checked prior to publication [Kraemer et al., 2015b]. The GBIF and Kraemer et al. data were cleaned to remove duplicate records, and reprojected to the global

equal area grid. To reduce spatial autocorrelation effects, we created a 5x5 km grid and only included one randomly selected record from grid cells with more than one record [Segurado et al., 2006, Hawkins, 2012].

To assess the relative importance of climate, habitat and resource constraints on niche use, we developed satellite-derived metrics that capture the spatial and temporal variation in temperature, precipitation, land cover and population densities. The spatial and temporal scales over which these patterns vary are quite distinct. While land cover can vary greatly over small spatial scales, it typically varies only minimally within a year in undisturbed landscapes. If land cover does change within a year, it is often the result of a transformative process (e.g., the conversion of forest to pasture). In contrast, temperature and precipitation can vary greatly on daily and monthly time scales, but at any time the spatial turnover in these patterns can be small relative to turnover in land cover. The derived covariates attempted to represent these spatial and temporal dynamics across scales for these niche constraints.

For temperature and precipitation, we derived mean, variance, and skewness statistics on a per-grid cell basis using all daily MODIS land surface temperature (LST; $°C$) and hourly TRMM precipitation measurements (PCP; $mm\,mo^{-1}$) from 2002-2017 [Justice et al., 1998, Huffman et al., 2007, Hou et al., 2013]. We calculated mean temperature and precipitation metrics to identify preferences for average conditions over a year, variance metrics to identify preference for static or dynamic climatic conditions, and skewness to calculate preference for anomalously hot/cold or wet/dry conditions [Huffman et al., 2007, Hou et al., 2013]. We included three land cover metrics to characterize habitat constraints: tree cover (TC; %) and the mean and variance of leaf area index (LAI; $m^2\,m^{-2}$) from 2002-2017 [Justice et al., 1998, Hansen et al., 2013]. The LAI covariates were selected to describe vegetation growth and phenology, as well as access to sugar feeding resources [Martinez-Ibarra et al., 1997, Chen and Kearney, 2015]. Tree cover was included to distinguish forests from agriculture and because each vector uses trees in urban and agricultural landscapes as habitat [Troyo et al., 2009, Landau and van Leeuwen, 2012]. To characterize resource constraints (i.e., available blood meal), we analyzed two population density covariates, human population density (POP; $people\,km^{-2}$) and livestock density (LIV; $animals\,km^{-2}$) from WorldPop and the Gridded Livestock of the World database [Tatem, 2017, Gilbert et al., 2018]. Human population density was $\log_{10}$ transformed to increase normality prior to analysis.

4.3.2 Species Distribution Modeling

To address our first research question, we evaluated niche preferences using the Maxent species distribution modeling software [Phillips et al., 2006, Steven J. Phillips, Miroslav Dudík, Robert E. Schapire, 2017], which uses logistic regression to calculate habitat suitability based on the conditional density of features at occurrence points and the unconditional density of features across the study area [Elith et al., 2011]. Prior to

analysis, we used the target group selection method [Phillips et al., 2009, Merow et al., 2013] to generate

a sampling bias adjustment using non-*Aedes* occurrence data from the Culicidae family and urbanization

data from nighttime lights [Mills et al., 2013]. We used these data to train a Maxent model, then used the

cumulative output, which indicates sampling frequency in human populations, as our sampling bias for all

other Maxent analyses. We therefore assumed that neither the mosquito vectors nor the vector sampling

efforts were uniformly distributed, but biased towards urban centers, so suitability was calculated relative to

a statistically-derived null distribution of areas where people live.

Maxent calculates relative species occurrence probabilities by comparing the statistical distributions of

environmental covariates at occurrence sites to a similar covariate distribution across potentially accessible

habitats (i.e., the background) per Eqn. 4.1:

$$Pr(y = 1|z) = \frac{f_1(z) \cdot Pr(y = 1)}{f(z)} \tag{4.1}$$

Where y represents a species, $y = 1$ are locations where that species was observed, z is a vector of

environmental covariates, $f(z)$ is a probability distribution of non-linear feature transformations derived from

the vector of covariates across the background, $f_1(z)$ is a probability distribution of features derived from the

covariates at species occurrence locations, and $Pr(y = 1|z)$ is the probability that a species occurs at a point

on the landscape as conditioned by the environment. The sampling bias adjustment modified the locations

from which the distribution $f(z)$ was drawn.

To identify which environmental covariates independently best predicted *Aedes* niche use, we ran four

sets of Maxent models, each using just the covariates from each potential niche axis (i.e., temperature,

precipitation, resource availability, and habitat). For population density we set $z = [log10(POP), LIV]$,

for temperature we set $z = [LST_{mean}, LST_{var}, LST_{skew}]$, for land cover we set $z = [LAI_{mean}, LAI_{var}, TC]$,

and for precipitation we set $z = [PCP_{mean}, PCP_{var}, PCP_{skew}]$. We evaluated model performance using the

area under the receiver operator curve (AUC), a metric of separability, which calculates the probability

that model predictions distinguish between suitable habitat and a semi-random geographic sample (i.e., the

background). We assessed model performance using 4-fold cross validation and reported a modified AUC

calculation (described in more detail below).

To assess the relative importance of each environmental covariate, we ran Maxent models using all covari-

ates, setting $z = [log10(POP), LIV, LST_{mean}, LST_{var}, LST_{skew}, LAI_{mean}, LAI_{var}, TC, PCP_{mean}, PCP_{var},$

$PCP_{skew}]$, and reported permutation importance scores. During model fitting, Maxent calculates how model

performance changes as feature coefficients change [Phillips and Dudík, 2008]. Permutation importance scores

are calculated by randomly altering the values of a single covariate and recalculating model performance

(here, AUC). These values are rescaled to a percentage based on how model performance changed for each

covariate permutation. The mean cross-validation permutation scores are reported in Table B.2.

Maxent models were run with the following parameters. For feature selection we ran models with only 'hinge' features enabled, which construct smooth, nonlinear response curves akin to a generalized additive model [Hastie and Tibshirani, 2004, Elith et al., 2011]. This was based on the assumption that *Aedes* vectors would respond in a continuous, nonlinear fashion to climate, habitat and resource patterns, akin to other temperature-dependent models that assume unimodal response functions [Mordecai et al., 2017]. To reduce model overfitting and to reduce the effects of collinearity in our environmental covariates, we increased the default β regularization parameter to 1.5, which penalizes complex models and shrinks model coefficients during training [Merow et al., 2013]. We selected this value by minimizing the difference between training and testing AUC scores in cross-validation using models trained with all covariates. For background selection, we preferentially sampled 10,000 random points using the sampling bias map described above. For the output format, we report Maxent's cumulative suitability metric, which is a relative occurrence rate rescaled between 0 and 100. This output can be interpreted as an omission rate; setting a threshold at a value of 10 to predict presence/absence will omit approximately 10% of presence records [Phillips and Dudík, 2008, Merow et al., 2013].

In our results we report a modified AUC score. AUC can be interpreted as the probability that a randomly chosen presence sample is ranked above a randomly chosen background sample. These probabilities are typically assessed using all test samples for presence and background points. However, when the number of background samples is greater than the number of presence samples, AUC values may appear artificially high by predicting low suitability across a large number of background samples without actually calculating appropriately high suitability scores in the small number of locations where a species is present. This is akin to a zero-inflation effect. We reduced this inflation effect by randomly selecting the same number of background samples as presence test samples (i.e., balancing the test dataset) before calculating AUC scores. This process was bootstrapped 5 times for each model to get an unbiased estimate of the adjusted AUC value, and we reported the bootstrapped mean.

4.3.3 Field Data Collection

To address our second research question comparing model results to field observations, we collected mosquito abundance data using an array of trapping methods across two regions, which included 38 field sites in Costa Rica and 96 field sites in Perú (Fig. 4.2; Table B.3). Both regions were sampled during the wet season—July 2017 in Costa Rica and November 2018 in Perú—when *Aedes* vectors and local mosquito surveillance programs were both active. The Costa Rica sites were located in the tropical southeast of the country near the border with Panamá, then considered a southern frontier for *Ae. albopictus*, which hadn't
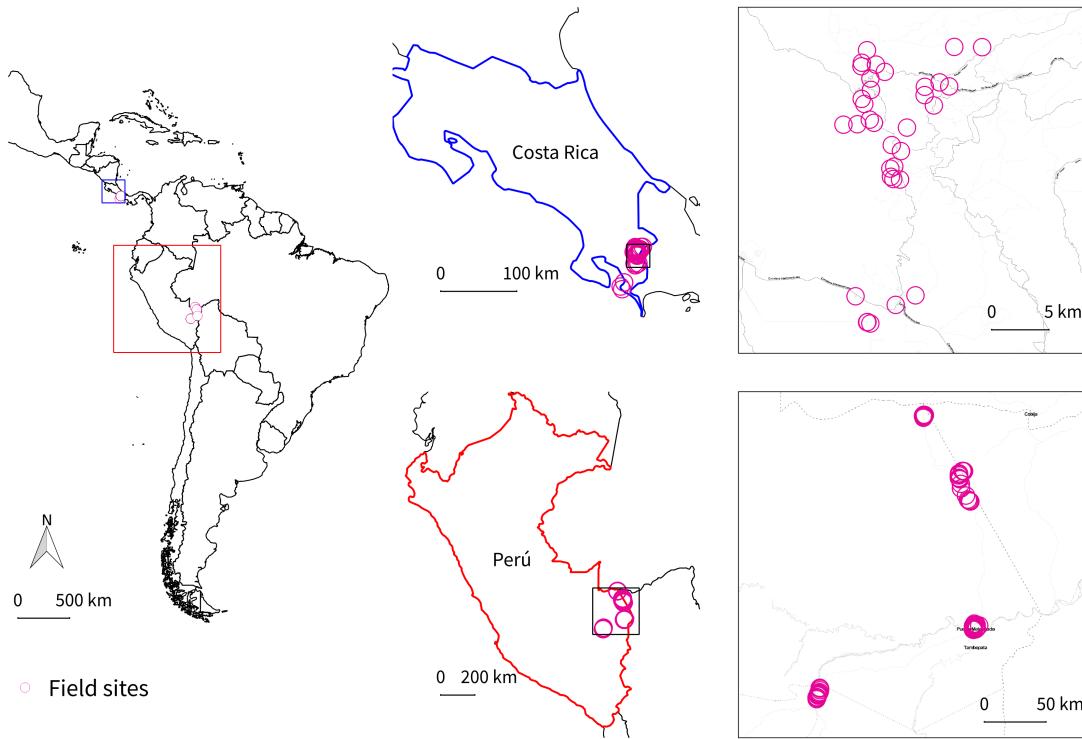
Figure 4.2: Field plot locations in Perú and Costa Rica. 134 field sites were visited, 38 in Costa Rica (top) and 96 in Perú, where four methods of mosquito trapping were used. These included BG-Sentinel traps, CDC light traps, Gravid *Aedes* traps (GAT), and aspiration. All sites were visited twice, with the exception of four sites in Costa Rica, which were visited once.

yet been formally reported in the region despite a recent increase in dengue cases [Gutiérrez, 2015]. The Perú sites were all located in the southeastern department of Madre de Dios. These included a set of sites near the border of Brazil and Bolivia, where there is concern that *Ae. albopictus* could migrate into Perú with traffic along the Interoceanic Highway. *Ae. albopictus* was the predominant vector of interest in the Costa Rica sites, with just three *Ae. aegypti* individuals identified there. *Ae. aegypti* was the only vector of interest identified in the Perú sites.

The site sampling scheme was designed to understand how *Aedes* habitat preferences and abundance patterns change along a land cover gradient. To identify sampling locations along this gradient, we created a four-class land cover map using k-means clustering. k-means is an unsupervised classification algorithm that iteratively seeds k cluster centroids in multivariate data and groups points according to the closest centroid, searching for centroids that minimize within-cluster variance. Generated using land cover covariates, these four classes roughly corresponded to forest, forest edge, agriculture and urban classes. Though logistical

challenges limited the number of forest and forest edge sites we could sample, it was important to trap in these areas. There is still active debate over the relationship between deforestation and mosquito-borne disease transmission [Norris, 2004, Tucker Lima et al., 2017], and few studies sample mosquito vector populations in forests prior to deforestation. Likewise, few SDMs include presence/absence test data in forests, making it difficult to evaluate the full extent of potential niche shifts forecast under global change.

Within the four land cover classes, we sampled opportunistically on properties where we could gather landowner approval, and landowners were notified if a vector was identified on their property. In urban areas, this included high and low density populations, including the rural mining town of Mazuko, Perú, and mid-elevation San Vito, Costa Rica, the capital of the Coto Brus district. Puerto Maldonado, Perú, the capital of Madre de Dios, was the most populous site sampled. The agriculture sites included cattle pastures as well as coffee, pineapple, banana and palm plantations. The forest and forest edge classes were grouped together, as trap placement was always $< 100 \ m$ from the forest edge, and because it was difficult to access many forested regions. These sites were located inside forest patches adjacent to agricultural areas, urban areas, and water bodies. In total, we sampled 36 forest/forest edge sites, 46 agriculture sites and 52 urban sites.

We sampled mosquito populations using four trapping methods: BG-Sentinel traps (baited with lures), CDC light traps (baited with CO_2 and octanol), Gravid Aedes traps (GATs), and manual aspiration. Traps were placed in the evening and picked up the following morning, remaining on site for approximately 12-16 hour periods. Aspiration was performed in the morning at the Costa Rica sites, and both in the evening and in the morning at the Perú sites. All sites were visited twice, with the exception of four Costa Rica sites that were visited once (PV-01, PV-02, PV-04, PV-05). Following trap retrieval and aspiration, all individuals were transferred to the lab for identification by trained personnel (MEH in Costa Rica, MSG and DN in Perú). Only high confidence identifications were flagged as positive vector observations; mature individuals and larval samples labeled as *Aedes* spp. were not included in the vector counts reported. In total, we collected 6,965 individual mosquitoes, 199 of which were identified as *Ae. aegypti* or *Ae. albopictus*.

4.3.4 Spatial Cross-Validation

To address our third research question, we evaluated evidence for niche conservation in each mosquito vector. In other *Aedes* modeling efforts, this has been done by evaluating model transferability: training SDMs in one region and evaluating how well model predictions transfer to geographically independent regions [Yates et al., 2018]. Models that fail to predict occurrence patterns outside the training region have been interpreted to indicate that these vectors are adapting to new climates; that niche shifts are underway, perhaps driven by hidden niche plasticity [Medley, 2010, Carlson et al., 2016]. This would fundamentally hinder our ability to
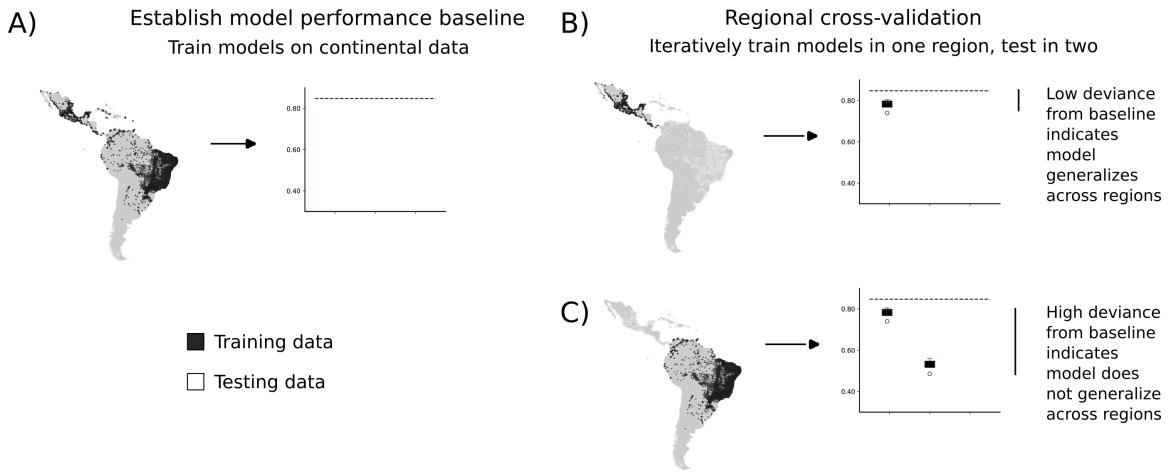
Figure 4.3: Methods summary for estimating niche conservation. A) First we used the species distribution models trained on each set of covariates (per *Section 4.3.2*) to calculate baseline model performance using all occurrence data. B) Then we iteratively trained models using just occurrence records from one of three regions— Mesoamerica, the Caribbean or South America—and tested performance on the remaining two regions. C) Similarity between baseline results and test results indicates that models generalize across regions, providing evidence of niche conservation.

forecast future distributions. However, a few key questions need to be addressed before the results of a model transferability analysis can be interpreted as evidence for or against niche evolution [Liu et al., 2020]. These include questions of spatial scale (do the occurrence records span the full extent of covariate space the vector occupies?), sample size (are there enough occurrence records to characterize vector-covariate relationships?) and representation (how well do the environmental covariates selected describe the different dimensions of the vector niche?).

We performed a spatial cross-validation analysis to address these questions. For each vector, the occurrence records were split by geographic region: into records from Mesoamerica, South America and the Caribbean. To quantify the effects of spatial scale, occurrence records from one region were used as training data and the remaining occurrence records were used as test data to evaluate each model. This process was repeated across each region (Fig. 4.3). The data were further subsampled within each region via 5-fold cross-validation to evaluate regional model variance and the effects of sample size.

To evaluate covariate representation, each model was trained using just covariates from each niche axis (resource availability, temperature, habitat, and precipitation). These models therefore evaluated how well a model trained on a subset of occurrence records from one region transfers to the remaining regions on a per-niche axis basis. We used the models trained across the whole study region as our "baseline" model performance for each niche axis. We interpret regional models that approximate baseline performance as

evidence of model transferability and therefore as evidence for niche conservatism. The reciprocal is not necessarily true, however. Models that deviate from the baseline are not necessarily evidence of niche evolution; these effects could be driven by issues of spatial scale, sample size or representation.
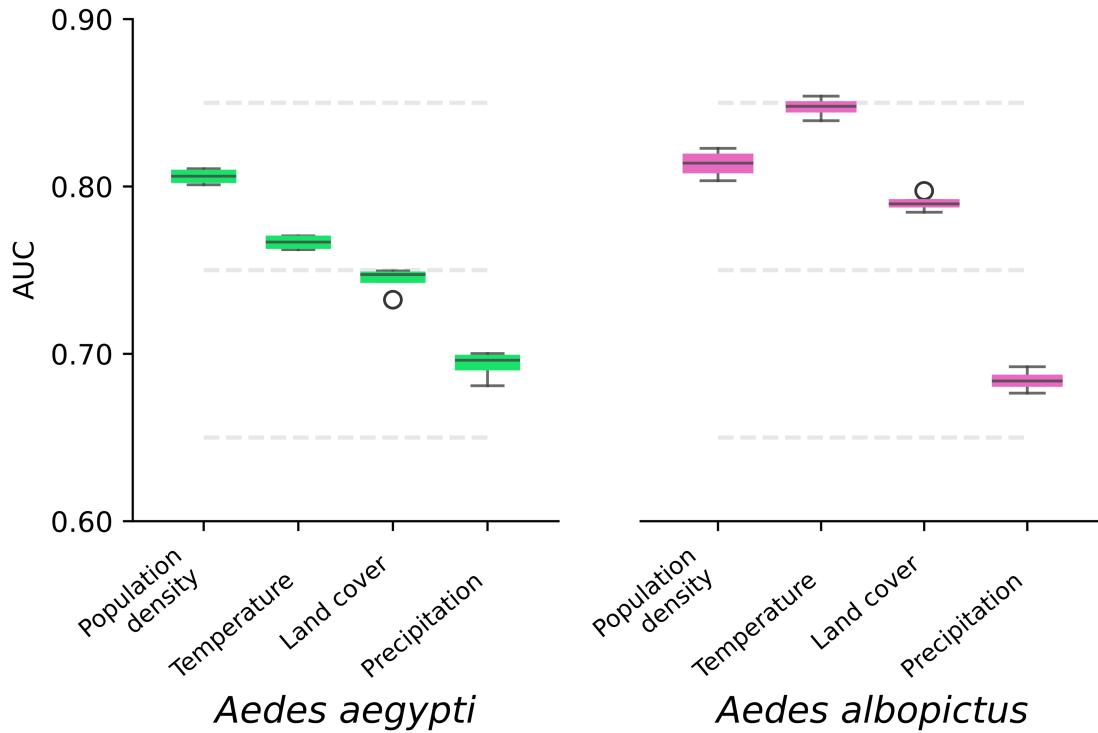
Figure 4.4: Population density and temperature alone predict *Aedes aegypti* and *Ae. albopictus* occurrence with high precision. These boxplots compare model performance among species distribution models that were trained on distinct covariate groups. Each set of models was trained only on covariates related to resource constraints (human population density, livestock density), temperature constraints (mean, variance and skewness of daily temperatures), habitat constraints (mean and variance of daily leaf area index, tree cover) or precipitation constraints (mean, variance and skewness of monthly rainfall). Uncertainty estimates were derived from 4-fold jackknifed cross-validation.

4.4 Results

4.4.1 Climate, Habitat and Resource Constraints

In models that isolated the predictive power of each of the three niche axes alone, we found that resource constraints, characterized by human population density and livestock density, best predicted the realized
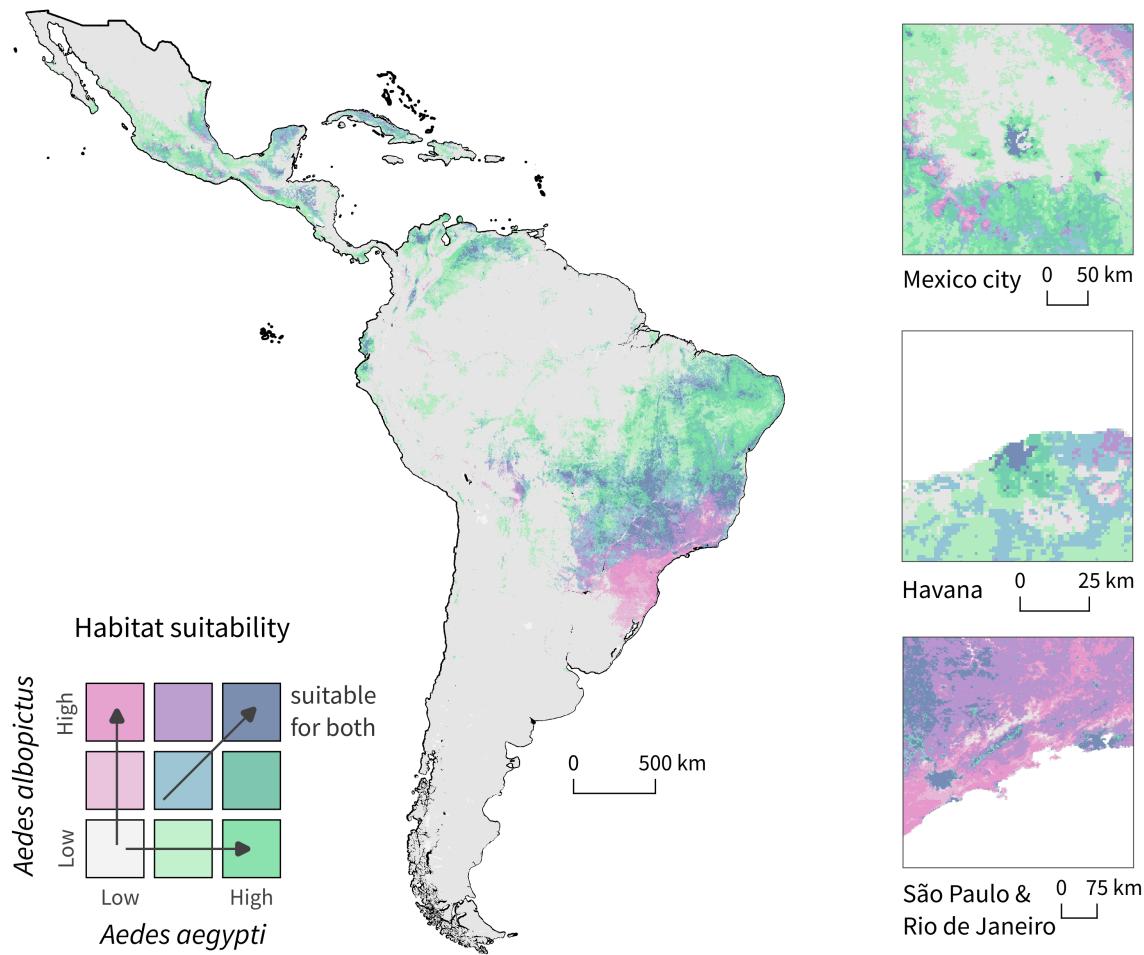
Figure 4.5: Fundamental niche estimates for each species estimated from the joint model trained on all 11 covariates show overlapping yet distinct fundamental niches for *Ae. aegypti* and *Ae. albopictus* across Latin America and the Caribbean. Insets show niche extents around population centers in Mesoamerica, the Caribbean and South America.

niche for *Ae. aegypti* (AUC mean = 0.806 ± 0.004 SD), while temperature constraints best predicted niche patterns for *Ae. albopictus* (AUC mean = 0.847 ± 0.005 SD; Fig. 4.4). For *Ae. aegypti*, resource constraints were followed by temperature (AUC mean = 0.767 ± 0.004 SD), land cover (AUC mean = 0.744 ± 0.007 SD), and precipitation constraints (AUC mean = 0.693 ± 0.007 SD) in discriminatory power. For *Ae. albopictus*, temperature constraints were followed by resource (AUC mean = 0.814 ± 0.007 SD), land cover (AUC mean = 0.790 ± 0.005 SD), and precipitation constraints (AUC mean = 0.684 ± 0.006 SD) in discriminatory power.

To assess the relative importance of each covariate, and to partially disentangle interactions between drivers, we trained models using all 11 environmental covariates and calculated permutation-based variable

importance scores (Table B.2) [Phillips and Dudík, 2008]. Human population density was the most important covariate for predicting *Ae. aegypti* distributions (17.9% of explained permutation variance), and daily temperature variation was the most important covariate for predicting *Ae. albopictus* distributions (51.9% explained permutation variance). While *Ae. aegypti* models were sensitive to a range of covariates (6 covariates each explained >10% of permutation variance), *Ae. albopictus* models relied on fewer covariates (only 3 covariates each explained >10% of permutation variance). These joint models, trained with all 11 covariates, performed better than any of the single niche-axis models reported above (4-fold cross-validation AUC mean = 0.836 ± 0.006 SD for *Ae. aegypti*, mean = 0.888 ± 0.005 SD for *Ae. albopictus*).

The predicted fundamental niche (i.e., potentially suitable habitat) for each vector is widely distributed throughout Central and South America and the Caribbean, particularly in coastal and lowland regions (Fig. 4.5). While many regions are predicted to be suitable for both vectors (Fig. 4.5, blue), distinct regions of high suitability for *Ae. aegypti* (green) or *Ae. albopictus* (pink) were interspersed at small spatial scales (e.g., Havana and Mexico City, Fig. 4.5 insets), as well as segregated across larger geographic regions (e.g., São Paulo and Rio de Janeiro for *Ae. albopictus*; the Atlantic dry forest region of Brazil and coastal and lowland regions of Mesoamerica for *Ae. aegypti*; 4.5).

4.4.2   Comparison to Lab and Field Observations

Both vectors occupied a narrow range of mean daily temperatures (Fig. 4.6), with *Ae. aegypti* occurring in warmer areas (95th percentile range 24.9-37.2°C, mean = 31.0°C) than *Ae. albopictus* (95th percentile range 22.1-35.9°C, mean = 29.1 °C; signed rank test $P < 0.001$), though *Ae. albopictus* mean daily temperature observations were not significantly different from the background (signed rank test $P = 0.430$). These mean observed daily temperature values are slightly higher than the thermal optima for dengue transmission predicted from a mechanistic model based on laboratory data (29.1°C for *Ae. aegypti*, 26.4°C for *Ae. albopictus*), within the confidence intervals of the thermal optima for fecundity (29.6°C for Ae. aegypti, 29.4°C for *Ae. albopictus*, and slightly cooler than the predicted optima for biting rates (33.8°C for *Ae. aegypti*, 31.8°C for *Ae. albopictus* (all based on trait thermal performance data summarized in [Mordecai et al., 2017]). Both vectors were observed in areas with significantly lower variance and higher skewness in daily temperatures than the background (all signed rank tests $P < 0.001$), indicating niche preferences for stable thermal conditions skewed towards warm extremes. All significance tests for stochastic equality were performed using the two-sided, nonparametric Brunner-Munzel test [Brunner and Munzel, 2000].

Mosquito abundance data from Perú and Costa Rica were grouped as forest/forest edge (36 sites), agriculture (46 sites) or urban (52 sites) to evaluate niche preferences and abundance patterns along a land use gradient. *Aedes* individuals were more abundant in urban sites than in agriculture or forest/forest edge
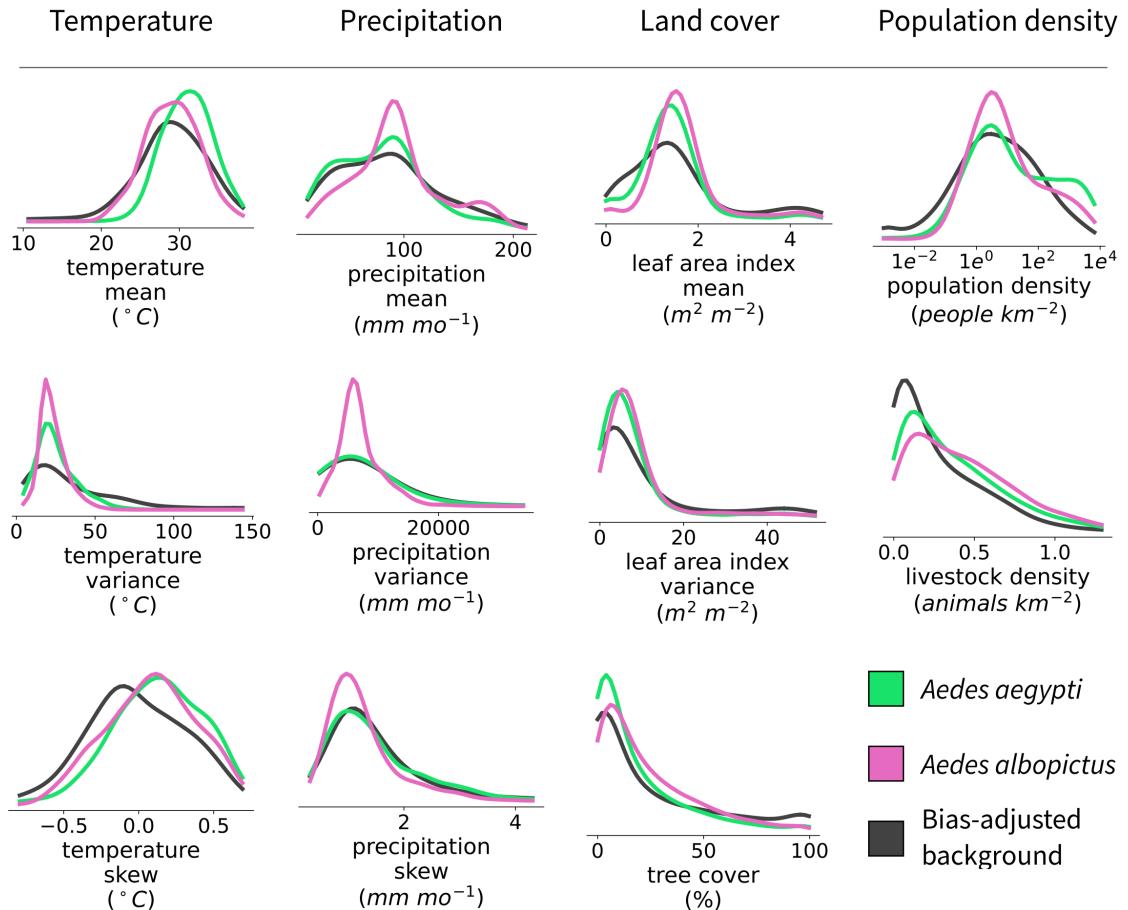
Figure 4.6: *Ae. aegypti* (green) and *Ae. albopictus* (pink) show distinct temperature, precipitation, habitat, and resource use profiles. These density distribution plots for 11 environmental covariates were extracted from occurrence points for each mosquito species and compared to a bias-adjusted sample of background points across Latin America and the Caribbean (grey). Differences between occurrence and background points indicate niche preferences for each species.

sites (Table B.3). We identified 156 out of 1,860 mosquitoes collected in urban sites as *Ae. aegypti* or *Ae. albopictus* (8.4% of individuals), occupying 76.5% and 62.9% of urban sites in Costa Rica and Perú, respectively. In agricultural sites, 34 out of 3,128 mosquitoes collected (1.1% of individuals) were identified as *Ae aegypti* or *Ae. albopictus*. Occupancy rates in agriculture sites varied by country, with 50% of sites occupied by *Ae. albopictus* in Costa Rica and 8.8% of sites occupied by *Ae aegypti* in Perú. We identified 9 out of 1,977 mosquitoes collected in forest/forest edge sites as *Ae aegypti* or *Ae. albopictus* (0.5% of individuals), occupying 27.3% of forest/forest edge sites in Perú but none in Costa Rica. *Ae aegypti* was the only vector of interest identified in the Perú sites and *Ae. albopictus* was the predominant vector of interest

identified in the Costa Rica sites.
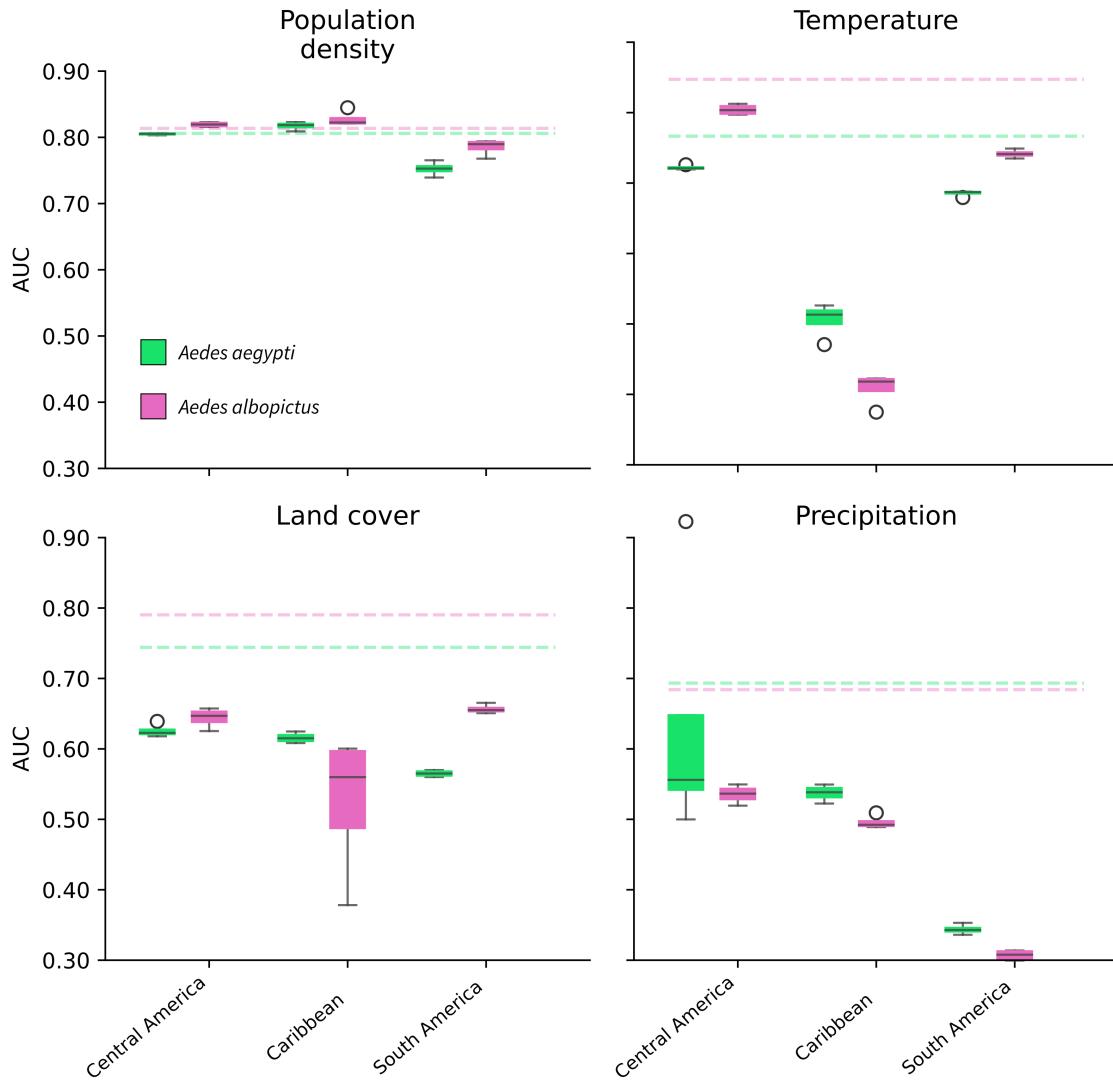
Figure 4.7: Spatial cross-validation analysis shows that models trained on just resource use covariates—human population density and livestock density—generalize across regions for both *Ae. aegypti* (green) and *Ae. albopictus* (pink; top left panel). By contrast, models trained on temperature (top right), land cover (bottom left), and precipitation (bottom right) on individual regions did not generalize to the remaining areas.

These field results corroborate the SDM results, where both vectors were observed in areas with dense human and livestock populations. *Ae. aegypti* was observed in areas with higher human population densities (mean = 632.2 *people km*$^{-2}$) than *Ae. albopictus* (mean = 363.6 *people km*$^{-2}$), both of which were higher

than the background (mean $= 157.8 \, people \, km^{-2}$, signed rank tests $P < 0.001$). *Ae. albopictus* was found in areas with higher livestock densities (mean $= 4.4 \, animals \, km^{-2}$) than the background (mean $= 2.8$ $animals \, km^{-2}$, signed rank test $P < 0.001$). Regarding habitat use, mean leaf area index patterns for *Ae. aegypti* were not significantly different from the background (signed rank test $P = 0.961$), but mean leaf area index and tree cover patterns for *Ae. albopictus* were significantly different from the background (signed rank test $P < 0.001$). Both vectors appear to avoid dense forests (95th percentile range 0.0-87.4% tree cover for *Ae. aegypti* and 0.1-86.1% tree cover for *Ae. albopictus*).

4.4.3 Niche Conservation

Models trained on just resource constraints within each region (Mesoamerica, the Caribbean, or South America) generalized well across other regions for both vectors (AUC mean range $= 0.752$-0.827 for both vectors in all regions, based on spatially-jackknifed cross-validation), suggesting consistent resource use patterns across the study area (Fig. 4.7). Temperature preferences generalized well for both vectors when trained on occurrence records from just Mesoamerica (AUC mean $= 0.722$ for *Ae. aegypti*, AUC mean $= 0.804$ for *Ae. albopictus*) or just South America (AUC mean $= 0.685$ for *Ae. aegypti*, AUC mean $= 0.742$ for *Ae. albopictus*), but did not generalize well for either vector when trained on records from just the Caribbean (AUC mean $= 0.506$ for *Ae. aegypti*, AUC mean $= 0.409$ for *Ae. albopictus*). Habitat use patterns did not generalize well across regions for either vector, though the degree to which these patterns did generalize was consistent across vectors and across regions (AUC mean range $= 0.524$-0.657). Precipitation patterns alone performed no better than random chance in most analyses (AUC mean range $= 0.307$-0.633).

4.5 Discussion

Mechanistically forecasting shifts in vector-borne disease burden with environmental change will be essential to manage and mitigate risks to exposed populations. While climatic constraints on *Aedes* distributions have been well characterized in the lab and by global spatial models, the mechanisms driving habitat and resource constraints were previously poorly characterized. Our results suggest that constraining forecasts based on habitat and resource availability is likely to significantly alter the geography of transmission under global change. We provide evidence that resource constraints (i.e., available blood meals) strongly predict *Aedes* distributions at continental scales, and that vector–resource relationships generalize across regions. Critically, resource constraints intersect with climate and habitat constraints to determine the two species' ranges, which are overlapping yet distinct. And while *Ae. aegypti* and *Ae. albopictus* are key arbovirus vectors—associated with urbanization, human-made breeding habitats, and human biting—few previous

niche modeling studies have explicitly included the density of humans and other blood meal resources as predictive covariates. With human-altered habitats expanding worldwide, the scope of invasion potential for these vectors is dramatic; previous work estimated that 80% of Brazil's population already lives in vector-suitable habitat [Cardoso-Leite et al., 2014]. And the gap between the fundamental and realized niches for these vectors is shrinking: the southward movement of *Ae. albopictus* has expanded to cover Costa Rica, where it was first officially recorded in northern provinces in the last decade [Calderón Arguedas et al., 2012, Marín Rodríguez et al., 2014].

Our approach differs from recent *Aedes* niche modeling studies in several regards. First, we exclusively used satellite-derived data as environmental covariates, which are measured continuously along regularly-spaced grids. This is a fundamentally different data collection strategy from interpolated weather station data, which are sparsely available in Latin America and the Caribbean [Fick and Hijmans, 2017] yet widely used in past studies. By aggregating daily satellite measurements over 16 years, we were able to construct rich, descriptive climate and habitat covariates that track large-scale spatiotemporal variation in each pattern. Second, our method for quantifying and adjusting for sampling bias significantly differs from other *Aedes* niche modeling studies. Since temperature patterns place a physiological limitation on where vectors can survive, researchers often constrain background sample selection to areas within an envelope of thermal viability [Brady et al., 2012, Bhatt et al., 2013, Kraemer et al., 2015a]. We instead assumed that, since vector sampling is a key part of epidemiological surveys, sampling bias was driven less by climatic constraints than by disease mitigation priorities in populated areas. We therefore quantified bias using urbanization data, prioritizing background selection from human-dominated areas, as background sampling locations should be selected with the same biases as the occurrence records [Phillips et al., 2009, Barbet-Massin et al., 2012, Fourcade et al., 2018]. Even after controlling for preferential sampling near human populations, we still found that population density and livestock density consistently predicted occurrence patterns, reinforcing the importance of resource constraints in driving vector distributions.

4.5.1 Mechanisms of Niche Use

Our data-driven results are supported by insights from mechanistic relationships between temperature, metabolism and transmission derived from lab tests for both vectors [Mordecai et al., 2017, Mordecai et al., 2019]. First, daily temperature patterns strongly predict niche use for both species, independently predicting occurrence patterns and driving model sensitivity in multivariate models (Fig. 4.4, Table B.2). Second, the density distributions of mean daily temperature are unimodal and peak between $29°C$ for *Ae. albopictus* and $31°C$ for *Ae. aegypti* (Fig. 4.6), which is similar to but slightly higher than the mechanistically predicted optima for dengue transmission of $26°C$ and $29°C$, respectively [Mordecai et al., 2017]. Third, *Ae. aegypti*

is more frequently observed at warmer temperatures than *Ae. albopictus*, tracking warmer thermal optima predicted for biting rates and immature survival rates. However, because these suitable temperatures are widespread in the tropics, including most of Latin America [Brady et al., 2012, Ryan et al., 2019], mean temperature alone is not a strong discriminant of *Aedes* occurrence. By contrast, temperature variance strongly predicted occurrence patterns, revealing niche preferences for a narrow envelope of thermal conditions amenable to year-round survival and biting.

Vegetation patterns better predicted the occurrence of *Ae. albopictus*, traditionally considered a more rural vector, than they did *Ae. aegypti*, typically an urban vector (Fig. 4.4). Within areas with suitable breeding habitats and access to blood meal hosts, vegetation provides more suitable microclimates, adult resting habitat, and nectar sources for sugar feeding to support metabolism [Marinotti et al., 1990, Martinez-Ibarra et al., 1997, Chen and Kearney, 2015]. Our vector surveys and other recent work [Troyo et al., 2009, Calderón-Arguedas et al., 2015] have identified agricultural areas and their adjacent forest elements as key breeding sites for *Ae. albopictus*, including coffee, palm, and pineapple plantations in Costa Rica, and for *Ae. aegypti*, as well as and areas outside mining camps in Perú. The species distribution models capture these niche preferences for areas with low to moderate tree cover and mean leaf area index, combined with high variance in leaf area index, an indicator of vegetation phenology and seasonal agricultural productivity (Fig. 4.6).

4.5.2 Mechanisms of Niche Conservation

We found strong evidence for niche conservation in resource use patterns via spatially-explicit cross-validation (Fig. 4.3). Models trained on just human population density and livestock density generalized well across regions, suggesting that these vectors have consistent resource requirements that explain spatial distribution patterns better than climate alone. *Ae. aegypti* occurred in higher population densities (mean $= 632$ $people\,km^{-2}$) than *Ae. albopictus* (mean $= 363$ $people\,km^{-2}$), both of which were higher than background values (mean $= 158$ $people\,km^{-2}$). This supports the characterization of *Ae. aegypti* as an "urban-affiliated" species (though it may be more precise to characterize them as "human-affiliated"). While *Ae. albopictus* also preferred human-dominated landscapes, sensitivity analysis showed that their occurrence records were better explained by livestock density (Table B.3). This is consistent with their weak but non-discriminating human biting habits [Gratz, 2004, Paupy et al., 2009, Kamgang et al., 2012]. Supplementing human blood meals with livestock blood meals is potentially an effective resource use strategy: livestock is the largest pool of mammal biomass worldwide, and their abundance and geographic footprint is increasing across the region [Bar-On et al., 2018, Nepstad et al., 2014a, Gilbert et al., 2018], especially as cattle ranching expands in Brazil under the current presidential administration [Rochedo et al., 2018, Kröger, 2020]. This generalist

resource use strategy may, in part, explain their invasion dynamics. The spread of *Ae. albopictus* across Latin America and the Caribbean, first reported in Texas and Brazil in the 1980s [Moore and Mitchell, 1997, Santos, 2003] —the epicenters of dispersal [Tatem et al., 2006, Wagman et al., 2013, Ortega-Morales and Rodríguez, 2016]—also coincided with the dramatic expansion of pastures across the region [Geist and Lambin, 2002, Barbier, 2004, Graesser et al., 2015].

While we did not see that niche preferences for temperature, land cover, or precipitation generalized across regions, which corroborates the results of similar studies [Medley, 2010, Pech-May et al., 2016, Carlson et al., 2016], we do not interpret this as evidence for niche evolution. Instead, we posit that vector–climate and vector–habitat relationships are complex and sensitive to a broad range of environmental variation, which are difficult to fully characterize over small geographic extents using statistically-driven SDMs. This may explain why temperature-driven models trained in the climatically diverse region of South America generalized well to Mesoamerica and the Caribbean, but models trained in the Caribbean were not reciprocal. We suggest that previous evidence for niche evolution may have been driven by this phenomenon and—critically— due to not accounting for niche conservation in resource use patterns, which did generalize. Resource use did generalize across the region, with a few occurrence records in the Caribbean predicting thousands more occurrence records across South America and Mesoamerica. Given these shortfalls, it is important to develop spatial modeling methods based on generalized, mechanistic vector–environment relationships. Our statistically-driven approach provided support for well-known vector–climate relationships, identified new vector–resource mechanisms driving distributions, and highlighted how the generality of vector–habitat relationships remains uncertain.

4.6   Conclusion

As both *Ae. aegypti* and *Ae. albopictus* prefer populated areas in warm climates with low to moderate vegetation cover, directional shifts under global environmental change are likely to expand the range of suitable vector habitat [Ryan et al., 2019]. We show here that it is critical to put people on the map in these global change scenarios [Ellis and Ramankutty, 2008], not only to characterize exposure and vulnerability to vector-borne disease but to better understand shifts in vector populations themselves and to design effective mitigation strategies. We also identified that cattle ranching may present an under-recognized risk driving *Ae. albopictus* invasion and potential arbovirus transmission, which is also forecast to expand in extent in the coming years. With dramatic global environmental changes rapidly approaching, public health planners should be prepared for a potentially devastating expansion of *Aedes* vectors and their arboviral passengers.

Chapter 5

Conclusion

5.1 New Horizons for Biodiversity Monitoring

It is not necessarily the case that access to more information on the state of biodiversity means that we will collectively make more informed choices about land use planning in favor of conservation. Dire and repeated warnings about the consequences of environmental change often appear unheeded, as local or national priorities—the scales at which all land management decisions are made—supersede global and ecological priorities [Ehrlich and Mooney, 1983, Daily, 1999, Barnosky et al., 2012, IPBES, 2019]. But information from monitoring efforts have led to dramatic changes in public policy in the past, like reducing the deforestation rate by 70% in the Amazon following some of the first satellite-based land cover change analyses [Roughgarden et al., 1991, Skole and Tucker, 1993, Nepstad et al., 2014b] and increasing reforestation in northwestern China following severe flooding and encroaching desertification [Fullen and Mitchell, 1994, Ye and Glantz, 2005]. Like independent journalism, biodiversity monitoring systems are a necessary but insufficient institution, informing public and official discourse on how human activities are changing global ecological processes and how these changes might be mitigated. Reporting this information clearly, credibly, and transparently will be especially important in the coming decade, as the greatest threat to conservation and climate change mitigation is likely to be the widespread transmission of bad-faith disinformation [Iyengar and Massey, 2018].

Satellites and other earth observing sensors are essential monitoring tools, providing repeat, thematically consistent, and globally available measurements of the earth's ecosystems. But these abundant measurements are difficult to translate into metrics that are biological, subject to change, and ecosystem agnostic, limiting adoption according to current biodiversity monitoring protocols [GEO BON, 2017]. This dissertation reviewed and addressed some of the technical and conceptual challenges to mapping biodiversity patterns from satellite

imagery, including addressing scale gaps between satellite and *in situ* data, using machine learning to simulate ecological processes, and handling biased or incomplete species data to characterize niche use patterns and predict spatial distributions. From these analyses emerges a flexible, scalable approach to measuring, monitoring and forecasting biodiversity change, highlighting opportunities to establish earth observations as the backbone of novel biodiversity monitoring systems. Three takeaways characterize the lessons learned from this work.

1. Biodiversity mapping analyses should include multi-scale environmental covariate data when linking satellite and *in situ* data. Local biodiversity patterns are often driven by ecological processes that operate at multiple spatial scales, and intermediate-scale patterns like disturbance regimes are often poorly represented in models that directly link fine-grained field data with coarse-grained environmental data. Multi-scale modeling approaches like [Baccini et al., 2017], reviewed in detail in *Chapter 2*, illustrate how integrating multiple data sources can translate biodiversity patterns from field to global scales, while characterizing the relative importance of multiple intersecting drivers of change.

2. Models can precisely translate earth observations measurements from units of energy into units of biodiversity when model form and covariate transformations are tailored to a specific domain of scale. Spectral reflectance patterns have different underlying drivers at the leaf, canopy, and community scales, and *Chapter 3* showed how isolating the drivers of canopy reflectance into discrete features dramatically improved species mapping accuracy, performing better than more complex neural network models trained with the same data [Marconi et al., 2018].

3. Biomimicry is a powerful approach to selecting and training machine learning algorithms to approximate and investigate ecological processes. While several processes can be characterized in lab environments, including temperature-dependent metabolic responses in ectotherms, many other processes like habitat use are harder to quantify in controlled settings. In *Chapter 4* we trained models that mimicked the form of a known ecological process, nonlinear thermal response functions, to fit similar functions quantifying mosquito-habitat and mosquito-resource relationships, which revealed previously under-recognized drivers of the spatial distributions of two globally invasive arbovirus disease vectors.

With more earth observations sensors slated to launch, access to biodiversity data increasing, and the development of new modeling approaches that integrate these data to produce globally-consistent metrics of biodiversity change, the promise of global biodiversity monitoring is becoming a reality. And it is not a moment too soon, with the effects of climate change and biodiversity loss already manifesting. How should we, collectively, use all of this new biodiversity information? And what purposes could monitoring systems serve if they were able to provide spatially and taxonomically complete information on the state of biodiversity?

Figure 5.1: A sign on Mount Chirripó in Costa Rica, home to the only high altitude Páramo grassland system in Mesoamerica, announcing that climate change has already arrived.

5.2 Acting on Complete Information

If the central challenge addressed in this dissertation was how to translate earth observations measurements from units of energy into units of biodiversity, the challenge posed by the preceding questions regards how to translate complete biodiversity information into effective conservation action. To clarify, by discussing "complete information" I'm encouraging a thought experiment regarding how we should use perfect knowledge of the state of the biosphere, and not suggesting that earth observations technologies alone will soon deliver this knowledge. This is certainly a topic too broad to address in much depth here. But as far as I can discern, there are four main opportunities for biodiversity monitoring systems to mitigate the risks posed by biodiversity change.

The first opportunity is to systematically quantify the conservation status of species, communities, and ecosystems to identify what is at risk, then design mitigation strategies. While this may seem a simplistic or

obvious task—and already the central goal of existing international conservation treaties like the Convention on Biological Diversity's Aichi Biodiversity Targets—existing efforts to mitigate risks are still severely data limited. [Geijzendorffer et al., 2016] found that, while global commitments to halting biodiversity loss have been signed, requiring extensive monitoring and evaluation to take action, the data available to evaluate conservation status in accordance with their reporting standards covers less than 25% of stated monitoring targets in some cases. This problem is especially acute for patterns of genetic diversity and ecosystem function, while data on trends in species populations are typically the most comprehensive. Incomplete information is a convenient and often legitimate excuse for inaction; complete information would remove this barrier.

The next opportunity is to analyze the historical satellite record and, based on the direction and magnitude of recent changes, design early-warning forecasts to predict upcoming threats to people and to species populations. As new data sharing policies have enabled open access to a long record of detailed satellite earth observations (Fig. 2.2), new temporally-explicit methods have been developed to map change over time using process-based and machine learning models [Cohen et al., 2018, Rao et al., 2020], which could be used to forecast near-term ecological changes and guide investments in conservation and restoration. A similar approach is being deployed in China following their recent National Ecosystem Assessment, where the impacts of a decade of land use change following a $50bn investment in environmental restoration and urbanization were used to guide a much larger investment in conservation in the coming years based on the trends that were declining [Ouyang et al., 2016, Bryan et al., 2018].

These kinds of large-scale investments aimed at improving the sustainability of land management practices have dramatically increased from both private and public sources over the past decade [Hamrick, 2016]. The monitoring, reporting, and verification standards associated with these investments are often inconsistently defined and applied, however, leaving many to wonder about the actual returns on these investments [Ferraro and Pattanayak, 2006, Engel et al., 2008, Sexton et al., 2016]. With so much at stake, there is a lot of pressure on scientists to quantitatively measure and monitor the changes that occur as a result these investments, and this emerged as a top priority following the previously mentioned National Ecosystem Assessment [Z. Ouyang, pers. communication]. The good news is that a global synthesis found the rate of biodiversity loss decreased directly in response the amount invested in conservation [Waldron et al., 2017]. The bad news is that the effectiveness of this spending decreased as development pressures increased, and over one third of the areas protected since 1992 have experienced and increase in development pressure [Jones et al., 2018]. Quantifying biodiversity change simultaneously with the drivers of change—putting people on the map [Ellis and Ramankutty, 2008]—is another key opportunity for monitoring systems.

The final opportunity, and the most urgent, is to increase public access to biodiversity information.

Access—the ability to derive benefits from resources [Ribot and Peluso, 2009]—to biodiversity itself is declining globally, due in part to increasing rates of urbanization [UN Population Division, 2019, Bratman et al., 2019] as well as the large-scale transfer of land from small private landowners and governments to large companies, often through land grabs [Peluso and Lund, 2011, Borras and Franco, 2012, Wolford et al., 2013]. Access to biodiversity information, as well as the technical expertise required to analyze and understand it, is especially limited in poor communities in disadvantaged regions—the communities being excluded from accessing nature's benefits—which is also where the negative effects of biodiversity and climate change are expected to be most severe (Fig. 5.1) [Borras et al., 2012, Turner, 2016, Barbier and Hochard, 2018, IPBES, 2019]. It is imperative that information detailing biodiversity change and the human impacts of these changes, such as shifting exposure to vector-borne diseases, be made both accessible and useful to the populations that are most vulnerable.

Complete information on the state of biodiversity alone is insufficient for catalyzing large-scale conservation action, and the degree to which we will collectively use such information to create a more just and sustainable world will depend on the ability and willingness of monitoring institutions to make biodiversity information clear, credible, and transparent to the public. And with intergovernmental organizations ceding the responsibility of monitoring to national and regional bodies, electing instead to focus on capacity building and filling data gaps [Larigauderie and Mooney, 2010, Scholes et al., 2012], perhaps there is an opportunity to build an independent global biodiversity monitoring institution based on these principles.

Appendix A

Glossary

- *Biodiversity pattern*: recurring and structured variation in the distributions of genes, species, communities and ecosystems.

- *Continuous measurements*: Earth observation measurements mapping the full geographical extent of a region without gaps.

- *Data dimensionality*: the minimum number of free variables needed to represent data without information loss [Camastra, 2003].

- *Discrete measurements*: Earth observation measurements mapping specific areas that do not cover the full geographical extent of a region.

- *Ecological processes*: Activities that result from interactions among organisms and between organisms and their environment[Martinez, 1996].

- *Earth observation sensor*: spaceborne or airborne instruments (e.g., a camera or radar) that record the electromagnetic radiation emitted or reflected by the landscape [Campbell and Wynne, 2011].

- *Extent*: the range over which a pattern or process occurs or is expected to occur [Nekola and White, 1999], such as a species fundamental niche, or the total area measured by an EO sensor.

- *Grain size*: the size of the smallest individual unit of measurement [Jensen and Lulla, 1987], such as a plot or transect in ecology, or the ground sampling distance of an Earth Observation sensor.

- *Multi-sensor fusion*: integrating measurements from multiple sensors with complementary spatial and temporal characteristics to characterise a single pattern [Hilker et al., 2009].

- *Radiometric calibration*: the conversion of raw image data (e.g. in digital number format) to units of

absolute radiance (e.g. in $W\,m^{-2}\,sr^{-1}\,\mu m^{-1}$) to standardise data from multiple sensors into a common scale [Chander et al., 2009].

- *Sensor fidelity*: the ability of a sensor to discriminate between land surface properties, and to discriminate signal from noise across the dynamic range of the sensor [Campbell and Wynne, 2011].

- *Sensor type*: general classifications of EO sensors based on the range of electromagnetic radiation measured, and how it was measured. Sensors are typically classified as active (i.e., sensors that emit their own energy, then record the reflection of that energy by the surface) or passive (i.e., sensors that measure energy emitted by the surface, not generated by the sensor). Radar sensors (e.g. Sentinel-1) are an example of active microwave (1 mm to 1 m) sensors. Multispectral sensors (e.g., Landsat) are an example of passive optical sensors that measure a range of typically visible (0.38–0.78 μm) to near-infrared (0.78–1.3 μm) or shortwave-infrared (1.3–3 μm) wavelengths [Campbell and Wynne, 2011].

Appendix B

Supplemental Figures and Tables
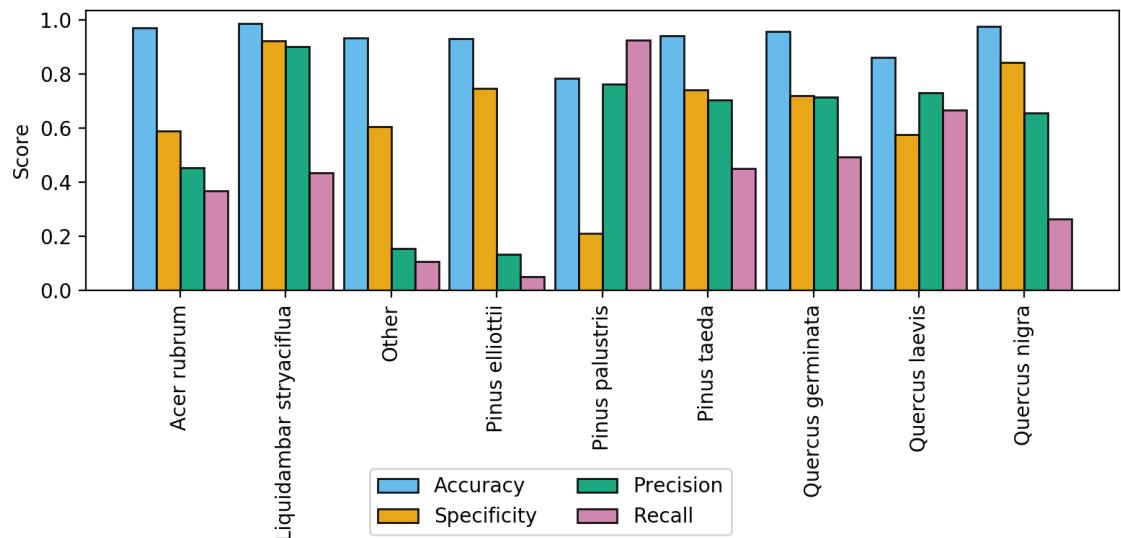
Figure B.1: Per-species secondary model performance metrics applied to test data calculated using per-crown prediction probabilities. Metrics weighted by the true negative rate (i.e., accuracy and specificity) were high for all species since the models correctly predicted the most common species. However, metrics weighted by the true positive rate (i.e., precision and recall) were more variable since there were fewer than six observed crowns for seven of the nine species. This penalized rare species misclassifications.

		Predicted								
	Species ID	*Acer rubrum*	*Liquidambar styraciflua*	Other	*Pinus elliottii*	*Pinus palustris*	*Pinus taeda*	*Quercus geminata*	*Quercus laevis*	*Quercus nigra*
Observed	*Acer rubrum*	**1**	0	0	0	0	0	0	0	1
	Liquidambar styraciflua	0	**1**	0	0	0	0	0	0	0
	Other	1	1	**1**	0	0	0	0	0	0
	Pinus elliottii	0	0	0	**0**	1	1	0	0	0
	Pinus palustris	0	0	0	2	**81**	0	0	1	0
	Pinus taeda	0	0	1	0	0	**4**	1	0	0
	Quercus geminata	0	0	0	0	0	0	**4**	0	0
	Quercus laevis	0	0	0	0	1	0	0	**22**	0
	Quercus nigra	0	0	0	0	0	0	0	0	**1**

Figure B.2: Confusion matrix computed from the binary classification results of the CCB-ID model on the competition test data. These metrics were calculated using the independent crown data. Bold entries highlight correct model predictions.

Sensor	EBV	Resoution (m)	Revisit time (d)	Sensor type
ALOS-AVNIR-2	Forest Cover	10	2	Multispectral
ALOS-PALSAR	Forest Cover	10-100	46	Radar-L-Band
ALOS-PALSAR-2	Forest Cover	3-60	46	Radar-L-Band
ASTER	Functional Composition	15-90	16	Multispectral-Thermal
AVHRR	Land Cover	1100	0.5	Multispectral-Thermal
CBERS-IRMSS/WFI	Land Cover	20-258	3-26	Multispectral
Cosmo-SkyMed-SAR	Aboveground Biomass	1-100	0.5-5	Radar-X-band
Envisat-AATSR	Disturbance Regime	1000	3	Multispectral-Thermal
Envisat-ASAR	Ecosystem Extent	30-1000	2-7	Radar-C-Band
ERS-ATSR	Disturbance Regime	1000	3	Multispectral-Thermal
ERS-AMI-SAR	Soil Moisture	30-1000	35	Radar-C-Band
Formosat-2	Land Cover	2-8	1	Multispectral
FY-1-MVISR	Disturbance Regime	1100	3-4	Multispectral-Thermal
FY-3-MERSI	Leaf Area Index	250-1100	1	Multispectral-Thermal
GeoEye-1	Species Occurrence	0.4-1.65	2.1-8.3	Multispectral
HJ-1A-HSI	Land Cover	100	4-31	Hyperspectral
IRMSS-HJ-1B	Disturbance Regime	150-300	4	Multispectral-Thermal
ICESat-GLAS	Tree Height	66	33-91	Large-footprint LiDAR
IKONOS	Taxonomic Diversity	1-4	3	Multispectral
Landsat-TM	Forest Cover	30-120	16	Multispectral-Thermal
Landsat-ETM+	Land Cover	30-60	16	Multispectral-Thermal
Landsat-OLI	Phenology	15-30	16	Multispectral-Thermal
EO-1-ALI	Land Cover	10-30	16	Multispectral
EO-1-Hyperion	Physiological Traits	30	16	Hyperspectral
MODIS	Leaf Area Index	250-1000	1-2	Multispectral-Thermal
Planet	Net primary productivity	3-5	1	Multispectral
Pleaides-HiRI	Taxonomic Diversity	0.7-2.8	2	Multispectral
QuickBird	Species Occurrence	2.4-0.6	2.5-5.6	Multispectral
RadarSat-1	Ecosystem Extent	8-100	24	Radar-C-Band
RadarSat-2	Ecosystem Extent	3-100	24	Radar-C-Band
RapidEye	Land Cover	5	1-5.5	Multispectral
Resourcesat-AWiFS	Disturbance Regime	5.8-56	5	Multispectral
RISAT-1	Taxonomic Diversity	3-50	25	Radar-C-Band
Sentinel-1A	Soil Moisture	5-25	12	Radar-C-Band
Sentinel-2A	Ecosystem Extent	10-20	5-10	Multispectral-Thermal
SPOT-HRV	Habitat Structure	10-20	2-3	Multispectral
SPOT-HRVIR	Ecosystem Extent	10-20	2-3	Multispectral
SPOT-Vegetation	Disturbance Regime	1000	1	Multispectral
SPOT-HRG	Forest Cover	2.5-20	2-3	Multispectral
Suomi NPP-VIIRS	Disturbance Regime	375-750	0.5-1	Multispectral-Thermal
Tandem-X	Aboveground Biomass	1-18	11	Radar-X-Band
TerraSAR-X	Functional Composition	1-18	11	Radar-X-Band
WorldView-2	Species Occurrence	0.5-1.8	1.1-3.7	Multispectral
WorldView-3	Species Occurrence	0.5-3.7	1-4.5	Multispectral

Table B.1: List of Satellite Earth observation sensors used to measure biodiversity patterns using the Essential Biodiversity Variables (EBV) framework.

Aedes aegypti			Aedes albopictus	
Population-log	17.9		LST-variance	51.9
LST-variance	13.3		Population-cattle	10.3
PCP-mean	12.3		PCP-variance	10
LAI-mean	12.1		PCP-mean	7.5
LAI-variance	11.4		PCP-skew	4.9
LST-mean	10.3		LST-mean	3.7
Population-cattle	9.5		Population-log	3.2
PCP-variance	8.2		LAI-mean	2.7
PCP-skew	2.8		LC-Trees	2.3
LST-skew	1.8		LST-skew	2
LC-Trees	0.3		LAI-variance	1.4

Table B.2: Sensitivity analysis showing permutation importance scores for 11 environmental features. Temperature variance is a strong independent predictor of *Aedes albopictus* spatial distributions, and *Ae. aegypti* was less sensitive to permutation in any one variable. These scores were generated by randomly permuting covariate values one by one, computing loss in model test performance, then scaling and ranking features based on the total range of model performance loss by vector.

Site Name	Country code	Latitude	Longitude	Aedes count	Non-Aedes count	Land cover class
SV01	CR	8.813	-82.967	0	3	urban
SV02	CR	8.834	-82.974	0	8	urban
SV05	CR	8.821	-82.967	1	6	urban
SV06	CR	8.827	-82.956	0	12	urban
ISLA	CR	8.832	-82.963	0	0	agriculture
SL01	CR	8.819	-82.912	5	15	urban
SL02	CR	8.815	-82.905	12	14	urban
CB01	CR	8.742	-82.944	3	2	urban
CB02	CR	8.744	-82.951	0	4	urban
CB03	CR	8.743	-82.949	14	23	urban
CB05	CR	8.751	-82.951	2	3	urban
CB04	CR	8.752	-82.948	2	18	urban
CN03	CR	8.650	-82.979	3	2	urban
OP01	CR	8.629	-82.967	1	10	agriculture
OP03	CR	8.630	-82.969	11	18	agriculture
OP03B	CR	8.630	-82.970	4	28	agriculture
CN04	CR	8.644	-82.947	1	209	urban
CN06	CR	8.651	-82.931	5	7	urban
SATE	CR	8.809	-82.924	2	2	agriculture
GABO	CR	8.802	-82.972	0	12	forest edge
GAPA	CR	8.806	-82.974	0	97	forest edge
RIJA	CR	8.787	-82.964	0	172	forest
MELI	CR	8.789	-82.968	0	244	forest
PV01	CR	8.427	-83.102	3	8	urban
PV02	CR	8.426	-83.101	0	2	agriculture
PV04	CR	8.379	-83.144	0	2	agriculture
PV05	CR	8.347	-83.126	0	1	forest
SAFR	CR	8.765	-82.943	0	7	agriculture
ELPU	CR	8.769	-82.950	0	5	agriculture
PINO	CR	8.844	-82.970	6	8	agriculture
SABO	CR	8.800	-82.917	5	2	urban
SV04	CR	8.831	-82.974	5	7	urban
FRAG	CR	8.785	-82.989	0	20	forest edge
FILA	CR	8.786	-82.978	0	16	forest
LOAN	CR	8.783	-82.938	0	204	forest
FELO	CR	8.846	-82.878	2	2	agriculture
VALE	CR	8.846	-82.900	0	7	agriculture
QUST	CR	8.815	-82.924	0	3	forest
PM01	PE	-12.589	-69.196	1	27	urban
PM02	PE	-12.596	-69.196	13	15	urban
PM03	PE	-12.598	-69.197	7	12	urban
PM04	PE	-12.594	-69.194	3	13	urban
PM05	PE	-12.593	-69.194	1	16	urban
PM06	PE	-12.588	-69.198	2	36	urban
PM07	PE	-12.589	-69.201	4	15	urban
PM08	PE	-12.594	-69.200	9	87	urban
PM09	PE	-12.605	-69.193	0	38	urban
PM10	PE	-12.584	-69.196	2	2	urban
PM11	PE	-12.584	-69.202	12	88	urban
PM12	PE	-12.578	-69.201	0	105	urban
PM13	PE	-12.577	-69.193	10	88	urban
PM14	PE	-12.572	-69.195	9	136	urban
PM15	PE	-12.562	-69.192	0	115	urban
PM16	PE	-12.571	-69.175	0	289	agriculture
PM17	PE	-12.570	-69.175	3	465	agriculture

Table B.3: Plot locations and abundance records from all *Aedes* and non-*Aedes* mosquito samples.

Site Name	Country code	Latitude	Longitude	Aedes count	Non-Aedes count	Land cover class
PM18	PE	-12.575	-69.180	0	131	urban
PM19	PE	-12.603	-69.199	5	20	urban
PM20	PE	-12.606	-69.202	3	79	agriculture
PM21	PE	-12.607	-69.197	1	14	urban
PM22	PE	-12.596	-69.173	0	49	agriculture
PM23	PE	-12.599	-69.180	0	2	agriculture
PM24	PE	-12.603	-69.181	2	208	agriculture
MA01	PE	-13.103	-70.368	0	4	urban
MA02	PE	-13.100	-70.367	1	27	urban
MA03	PE	-13.099	-70.372	0	4	urban
MA04	PE	-13.098	-70.369	2	5	urban
MA05	PE	-13.100	-70.367	1	4	urban
MA06	PE	-13.102	-70.368	8	1	urban
MA07	PE	-13.103	-70.370	0	8	urban
MA08	PE	-13.105	-70.369	1	7	urban
MA09	PE	-13.103	-70.369	1	10	urban
MA10	PE	-13.062	-70.350	0	5	forest edge
MA11	PE	-13.048	-70.346	0	2	agriculture
MA12	PE	-13.037	-70.345	0	28	forest
MA13	PE	-13.030	-70.348	0	10	forest
MA14	PE	-13.069	-70.354	0	11	forest
MA15	PE	-13.077	-70.359	0	221	urban
MA16	PE	-13.113	-70.372	0	14	urban
MA18	PE	-13.086	-70.362	0	11	urban
MA19	PE	-13.086	-70.363	0	31	urban
MA20	PE	-13.081	-70.360	0	2	urban
MA23	PE	-13.073	-70.358	1	22	urban
MA21	PE	-13.072	-70.358	1	16	urban
MA22	PE	-13.072	-70.360	0	61	agriculture
MA24	PE	-13.084	-70.362	0	16	urban
VP01	PE	-11.043	-69.573	0	48	agriculture
VP02	PE	-11.043	-69.573	0	23	agriculture
VP04	PE	-11.042	-69.574	0	173	agriculture
VP05	PE	-11.042	-69.572	0	40	agriculture
VP06	PE	-11.043	-69.574	0	3	agriculture
VP07	PE	-11.043	-69.575	0	17	agriculture
VP02	PE	-11.043	-69.573	0	23	agriculture
VP03	PE	-11.042	-69.571	0	11	agriculture
VP09	PE	-11.043	-69.574	0	4	agriculture
VP08	PE	-11.043	-69.574	0	19	agriculture
VP10	PE	-11.033	-69.568	0	3	agriculture
VP11	PE	-11.034	-69.570	0	10	agriculture
VP12	PE	-11.035	-69.571	0	13	agriculture
VP13	PE	-11.036	-69.572	0	35	agriculture
VP14	PE	-11.042	-69.574	0	186	agriculture
VP15	PE	-11.042	-69.575	0	29	agriculture
VP19	PE	-11.035	-69.562	0	19	agriculture
VP20	PE	-11.035	-69.563	0	2	forest edge
VP21	PE	-11.036	-69.564	0	9	forest edge
VP22	PE	-11.036	-69.563	0	7	agriculture
VP23	PE	-11.036	-69.563	0	19	agriculture
VP24	PE	-11.036	-69.563	0	59	agriculture
VP16	PE	-11.047	-69.574	0	5	agriculture
VP17	PE	-11.055	-69.572	0	17	agriculture
VP18	PE	-11.040	-69.576	0	13	agriculture

Site Name	Country code	Latitude	Longitude	Aedes count	Non-Aedes count	Land cover class
IB01	PE	-11.669	-69.221	2	70	forest
IB02	PE	-11.667	-69.226	0	38	forest
IB03	PE	-11.662	-69.234	0	12	agriculture
IB04	PE	-11.655	-69.240	0	118	agriculture
IB05	PE	-11.625	-69.254	0	11	forest edge
IB06	PE	-11.573	-69.291	0	31	forest edge
IB07	PE	-11.573	-69.292	0	21	forest
IB08	PE	-11.539	-69.293	0	114	forest
IB09	PE	-11.506	-69.300	0	130	forest edge
IB10	PE	-11.502	-69.302	0	133	forest edge
IB11	PE	-11.500	-69.303	1	59	forest edge
IB12	PE	-11.495	-69.308	1	91	forest edge
IB13	PE	-11.478	-69.308	0	44	forest edge
IB15	PE	-11.463	-69.306	2	87	forest edge
IB16	PE	-11.447	-69.268	0	35	forest
IB17	PE	-11.447	-69.270	0	105	forest
IB18	PE	-11.447	-69.270	0	15	forest
IB19	PE	-11.446	-69.272	0	18	forest
IB20	PE	-11.446	-69.273	1	27	forest
IB21	PE	-11.446	-69.274	2	39	forest
IB22	PE	-11.446	-69.276	0	32	forest
IB23	PE	-11.445	-69.279	0	29	forest
IB24	PE	-11.445	-69.281	0	5	forest
IB25	PE	-11.470	-69.305	0	259	agriculture
PM25	PE	-12.576	-69.155	0	706	agriculture

Appendix C

Other Published Work

Below is a list of peer-reviewed work I contributed to during the writing of this dissertation.

- Hendershot, J. N., Smith, J. R., Anderson, C. B., Letten, A. D., Frishkoff, L. O., Zook, J. R., Fukami, T. & Daily, G. C. (2020). Intensive farming drives long-term shifts in avian community composition. *Nature*, 579(7799), 393-396.

- Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., De Vries, S., Flanders, J., Folke, C., Frumkin, H., Gross, J. J., Hartig, T., Kahn, Jr., P. H., Kuo, M., Lawler, J. J., Levin, P. S., Lindahl, T., Meyer-Lindenberg, A., Mitchell, R., Ouyang, Z., Roe, J., Scarlett, L., Smith, J. R., van den Bosch, M., Wheeler, B. W., White, M., Zheng, H. & Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science advances*, 5(7), eaax0903.

- Ramirez-Reyes, C., Brauman, K. A., Chaplin-Kramer, R., Galford, G. L., Adamo, S. B., Anderson, C. B., Allington, G. R., Bagstad, K. J., Coe, M. T., Cord, A. F., Dee, L.E., Gould, R. K., Jain, M., Kowal, V. A., Muller-Karger, F., Norriss, J., Potapov, P., Qiu, J., Rieb, J. T.,, Robinson, B. E., Samberg, L. H., Singh, N., Szeto, S. H., Voigt, B., Watson, K. & Wright, T. M (2019). Reimagining the potential of Earth observations for ecosystem service assessments. *Science of the Total Environment*, 665, 1053-1063.

- Smith, J. R., Letten, A. D., Ke, P. J., Anderson, C. B., Hendershot, J. N., Dhami, M. K., Dlott, G. A., Grainger, T. N., Howard, M. E., Morrison, B. M. L., Routh, D., San Juan, P. A., Mooney, H. A., Mordecai, E. A., Crowther, T. W., & Daily, G. C. (2018). A global test of ecoregions. *Nature Ecology & Evolution*, 2(12), 1889-1896.

Appendix D

License Terms for Chapter 2 Reproduction

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Oct 24, 2020

Licensed Content Pages: 14

Type of use: Dissertation/Thesis

Requestor type: Author of this Wiley article

Format: Print and electronic

Portion: Full article

Will you be translating? No

Title: Biodiversity monitoring, earth observations and the ecology of scale

Institution name: Stanford University

Expected presentation date: Dec 2020

Requestor Location:

Christopher Anderson

557 Vermont St.

San Francisco, CA 94107

United States

Publisher Tax ID: EU826007151

Total: $0.00 USD

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at `http://myaccount.copyright.com`).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, and any CONTENT (PDF or image file) purchased as part of your order, is for a one-time use only and limited to any maximum distribution number specified in

the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts, You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED

IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-[commercial use].

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. (see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library `http://olabout.wiley.com/WileyCDA/Section/id-410895.html`

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Bibliography

[Adams et al., 2013] Adams, H. D., Williams, A. P., Xu, C., Rauscher, S. A., Jiang, X., and McDowell, N. G. (2013). Empirical and process-based approaches to climate-induced forest mortality models. *Frontiers in Plant Science*, 4:438.

[Adler et al., 2011] Adler, P. B., Seabloom, E. W., Borer, E. T., Hillebrand, H., Hautier, Y., Hector, A., Harpole, W. S., O'Halloran, L. R., Grace, J. B., Anderson, T. M., Bakker, J. D., Biederman, L. A., Brown, C. S., Buckley, Y. M., Calabrese, L. B., Chu, C.-J., Cleland, E. E., Collins, S. L., Cottingham, K. L., Crawley, M. J., Damschen, E. I., Davies, K. F., DeCrappeo, N. M., Fay, P. A., Firn, J., Frater, P., Gasarch, E. I., Gruner, D. S., Hagenah, N., Hille Ris Lambers, J., Humphries, H., Jin, V. L., Kay, A. D., Kirkman, K. P., Klein, J. A., Knops, J. M. H., La Pierre, K. J., Lambrinos, J. G., Li, W., MacDougall, A. S., McCulley, R. L., Melbourne, B. A., Mitchell, C. E., Moore, J. L., Morgan, J. W., Mortensen, B., Orrock, J. L., Prober, S. M., Pyke, D. A., Risch, A. C., Schuetz, M., Smith, M. D., Stevens, C. J., Sullivan, L. L., Wang, G., Wragg, P. D., Wright, J. P., and Yang, L. H. (2011). Productivity is a poor predictor of plant species richness. *Science*, 333(6050):1750–1753.

[Altizer et al., 2013] Altizer, S., Ostfeld, R. S., Johnson, P. T. J., Kutz, S., and Harvell, C. D. (2013). Climate change and infectious diseases: from evidence to a predictive framework. *Science*, 341(6145):514–519.

[Anderson, 2012] Anderson, L. O. (2012). Biome-Scale forest properties in amazonia based on field and satellite observations. *Remote Sensing*, 4(5):1245–1271.

[Arita and Rodriguez, 2002] Arita, H. T. and Rodriguez, P. (2002). Geographic range, turnover rate and the scaling of species diversity. *Ecography*, 25(5):541–550.

[Asner, 1998] Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sens. Environ.*, 64(3):234–253.

[Asner et al., 2016] Asner, G. P., Knapp, D. E., Anderson, C. B., Martin, R. E., and Vaughn, N. (2016). Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proc. Natl. Acad. Sci. U. S. A.*, 113(28):E4043–51.

[Asner et al., 2012] Asner, G. P., Knapp, D. E., Boardman, J., Green, R. O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R. E., Anderson, C., and Field, C. B. (2012). Carnegie airborne observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sens. Environ.*, 124(Supplement C):454–465.

[Asner and Martin, 2008] Asner, G. P. and Martin, R. E. (2008). Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sens. Environ.*, 112(10):3958–3970.

[Asner et al., 2015] Asner, G. P., Martin, R. E., Anderson, C. B., and Knapp, D. E. (2015). Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sens. Environ.*, 158:15–27.

[Asner et al., 2011] Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C., Carranza, L., Martinez, P., Houcheime, M., Sinca, F., and Weiss, P. (2011). Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sens. Environ.*, 115(12):3587–3598.

[Asner et al., 1998] Asner, G. P., Wessman, C. A., and Archer, S. (1998). Scale dependence of absorption of photosynthetically active radiation in terrestrial ecosystems. *Ecological Applications*, 8(4):1003–1021.

[Baccini et al., 2012] Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., and Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Chang.*, 2:182.

[Baccini et al., 2017] Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., and Houghton, R. A. (2017). Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, 358(6360):230–234.

[Bachmann et al., 2015] Bachmann, M., Makarau, A., Segl, K., and Richter, R. (2015). Estimating the influence of spectral and radiometric calibration uncertainties on EnMAP data Products—Examples for ground reflectance retrieval and vegetation indices. *Remote Sensing*, 7(8):10689–10714.

[Baldeck and Asner, 2014] Baldeck, C. A. and Asner, G. P. (2014). Improving remote species identification through efficient training data collection. *Remote Sensing*, 6(4):2682–2698.

[Baldeck et al., 2015] Baldeck, C. A., Asner, G. P., Martin, R. E., Anderson, C. B., Knapp, D. E., Kellner, J. R., and Wright, S. J. (2015). Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PLoS One*, 10(7):e0118403.

[Baldeck et al., 2014] Baldeck, C. A., Colgan, M. S., Féret, J. B., Levick, S. R., Martin, R. E., and Asner, G. P. (2014). Landscape-scale variation in plant community composition of an african savanna from airborne species mapping. *Ecol. Appl.*, 24(1):84–93.

[Bar-On et al., 2018] Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on earth. *Proc. Natl. Acad. Sci. U. S. A.*, 115(25):6506–6511.

[Barbet-Massin et al., 2012] Barbet-Massin, M., Jiguet, F., Albert, C. H., and Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol. Evol.*, 3(2):327–338.

[Barbier, 2004] Barbier, E. B. (2004). Agricultural expansion, resource booms and growth in latin america: Implications for long-run economic development. *World Dev.*, 32(1):137–157.

[Barbier and Hochard, 2018] Barbier, E. B. and Hochard, J. P. (2018). The impacts of climate change on the poor in disadvantaged regions. *Rev Environ Econ Policy*, 12(1):26–47.

[Barnosky et al., 2012] Barnosky, A. D., Hadly, E. A., Bascompte, J., Berlow, E. L., Brown, J. H., Fortelius, M., Getz, W. M., Harte, J., Hastings, A., Marquet, P. A., Martinez, N. D., Mooers, A., Roopnarine, P., Vermeij, G., Williams, J. W., Gillespie, R., Kitzes, J., Marshall, C., Matzke, N., Mindell, D. P., Revilla, E., and Smith, A. B. (2012). Approaching a state shift in earth's biosphere. *Nature*, 486(7401):52–58.

[Bartsch et al., 2009] Bartsch, A., Wagner, W., Scipal, K., Pathe, C., Sabel, D., and Wolski, P. (2009). Global monitoring of wetlands–the value of ENVISAT ASAR global mode. *J. Environ. Manage.*, 90(7):2226–2233.

[Basu et al., 2015] Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., and Nemani, R. (2015). Deepsat: a learning framework for satellite imagery. In *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems*, pages 1–10.

[Baumgartner et al., 2012] Baumgartner, A., Gege, P., Köhler, C., Lenhard, K., and Schwarzmaier, T. (2012). Characterisation methods for the hyperspectral sensor HySpex at DLR's calibration home base. In *SPIE Remote Sensing 2012*, pages 1–8. elib.dlr.de.

[Beck et al., 2014] Beck, J., Böller, M., Erhardt, A., and Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol. Inform.*, 19:10–15.

[Bhatt et al., 2013] Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., Myers, M. F., George, D. B., Jaenisch, T., Wint, G. R. W., Simmons, C. P., Scott, T. W., Farrar, J. J., and Hay, S. I. (2013). The global distribution and burden of dengue. *Nature*, 496(7446):504–507.

[Bini et al., 2006] Bini, L. M., Diniz-Filho, J. A. F., Rangel, T. F. L. V. B., Bastos, R. P., and Pinto, M. P. (2006). Challenging wallacean and linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Divers. Distrib.*, 12(5):475–482.

[Blaschke et al., 2008] Blaschke, T., Lang, S., and Hay, G. (2008). *Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications*. Springer Science & Business Media.

[Boakes et al., 2010] Boakes, E. H., McGowan, P. J. K., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K., and Mace, G. M. (2010). Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biol.*, 8(6):e1000385.

[Booth et al., 2014] Booth, T. H., Nix, H. A., Busby, J. R., and Hutchinson, M. F. (2014). bioclim : the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Divers. Distrib.*, 20(1):1–9.

[Borras et al., 2012] Borras, S. M., Franco, J. C., Gómez, S., Kay, C., and Spoor, M. (2012). Land grabbing in latin america and the caribbean. *J. Peasant Stud.*, 39(3-4):845–872.

[Borras and Franco, 2012] Borras, Jr, S. M. and Franco, J. C. (2012). Global land grabbing and trajectories of agrarian change: A preliminary analysis. *Journal of Agrarian Change*, 12(1):34–59.

[Bradley et al., 2007] Bradley, B. A., Jacob, R. W., Hermance, J. F., and Mustard, J. F. (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sens. Environ.*, 106(2):137–145.

[Brady et al., 2012] Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., Moyes, C. L., Farlow, A. W., Scott, T. W., and Hay, S. I. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl. Trop. Dis.*, 6(8):e1760.

[Brady et al., 2014] Brady, O. J., Golding, N., Pigott, D. M., Kraemer, M. U. G., Messina, J. P., Reiner, Jr, R. C., Scott, T. W., Smith, D. L., Gething, P. W., and Hay, S. I. (2014). Global temperature constraints on aedes aegypti and ae. albopictus persistence and competence for dengue virus transmission. *Parasit. Vectors*, 7:338.

[Brandt et al., 2017] Brandt, L. A., Benscoter, A. M., Harvey, R., Speroterra, C., Bucklin, D., Romañach, S. S., Watling, J. I., and Mazzotti, F. J. (2017). Comparison of climate envelope models developed using expert-selected variables versus statistical selection. *Ecol. Modell.*, 345(Supplement C):10–20.

[Bratman et al., 2019] Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., de Vries, S., Flanders, J., Folke, C., Frumkin, H., Gross, J. J., Hartig, T., Kahn, P. H., Kuo, M., Lawler, J. J., Levin, P. S., Lindahl, T., Meyer-Lindenberg, A., Mitchell, R., Ouyang, Z., Roe, J., Scarlett, L., Smith, J. R., van den Bosch, M., Wheeler, B. W., White, M. P., Zheng, H., and Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science Advances*, 5(7):eaax0903.

[Brazil, 2019] Brazil, U. S. M. (2019). Health alert: Dengue in brazil — U.S. embassy & consulates in brazil. `https://br.usembassy.gov/health-alert-dengue-in-brazil-november-7-2019/`. Accessed: 2020-3-17.

[Breiman, 2001] Breiman, L. (2001). Random forests. *Mach. Learn.*, 45(1):5–32.

[Brito, 2010] Brito, D. (2010). Overcoming the linnean shortfall: Data deficiency and biological survey priorities. *Basic Appl. Ecol.*, 11(8):709–713.

[Brodrick et al., 2019] Brodrick, P. G., Davies, A. B., and Asner, G. P. (2019). Uncovering ecological patterns with convolutional neural networks. *Trends Ecol. Evol.*

[Broge and Leblanc, 2001] Broge, N. H. and Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.*, 76(2):156–172.

[Brown et al., 2004] Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). TOWARD a METABOLIC THEORY OF ECOLOGY. *Ecology*, 85(7):1771–1789.

[Brown et al., 2006] Brown, M. E., Pinzon, J. E., Didan, K., Morisette, J. T., and Tucker, C. J. (2006). Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SeaWiFS, MODIS, and landsat ETM+ sensors. *IEEE Trans. Geosci. Remote Sens.*, 44(7):1787–1793.

[Brunner and Munzel, 2000] Brunner, E. and Munzel, U. (2000). The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 42(1):17–25.

[Bryan et al., 2018] Bryan, B. A., Gao, L., Ye, Y., Sun, X., Connor, J. D., Crossman, N. D., Stafford-Smith, M., Wu, J., He, C., Yu, D., Liu, Z., Li, A., Huang, Q., Ren, H., Deng, X., Zheng, H., Niu, J., Han, G., and Hou, X. (2018). China's response to a national land-system sustainability emergency. *Nature*, 559(7713):193–204.

[Butchart et al., 2010] Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., Galloway, J. N., Genovesi, P., Gregory, R. D., Hockings, M., Kapos, V., Lamarque, J.-F., Leverington, F., Loh, J., McGeoch, M. A., McRae, L., Minasyan, A., Hernández Morcillo, M., Oldfield, T. E. E., Pauly, D., Quader, S., Revenga, C., Sauer, J. R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S. N., Symes, A., Tierney, M., Tyrrell, T. D., Vié, J.-C., and Watson, R. (2010). Global biodiversity: indicators of recent declines. *Science*, 328(5982):1164–1168.

[Butler, 2014] Butler, D. (2014). Many eyes on earth. *Nature*, 505(7482):143–144.

[Calabrese et al., 2014] Calabrese, J. M., Certain, G., Kraan, C., and Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models: Stacking species distribution models. *Glob. Ecol. Biogeogr.*, 23(1):99–112.

[Calderón Arguedas et al., 2012] Calderón Arguedas, O., Troyo, A., Avendaño, A., and Gutiérrez, M. (2012). Aedes albopictus (skuse) en la región huetar atlántica de costa rica. *Revista Costarricense de Salud Pública*, 21(2):76–80.

[Calderón-Arguedas et al., 2015] Calderón-Arguedas, O., Troyo, A., Moreira-Soto, R. D., Marín, R., and Taylor, L. (2015). Dengue viruses in aedes albopictus skuse from a pineapple plantation in costa rica. *J. Vector Ecol.*, 40(1):184–186.

[Camastra, 2003] Camastra, F. (2003). Data dimensionality estimation methods: a survey. *Pattern Recognit.*, 36(12):2945–2954.

[Campbell and Wynne, 2011] Campbell, J. B. and Wynne, R. H. (2011). *Introduction to Remote Sensing, Fifth Edition*. Guilford Press.

[Campbell et al., 2015] Campbell, L. P., Luther, C., Moo-Llanes, D., Ramsey, J. M., Danis-Lozano, R., and Peterson, A. T. (2015). Climate change influences on global distributions of dengue and chikungunya virus vectors. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 370(1665).

[Cardoso-Leite et al., 2014] Cardoso-Leite, R., Vilarinho, A. C., Novaes, M. C., Tonetto, A. F., Vilardi, G. C., and Guillermo-Ferreira, R. (2014). Recent and future environmental suitability to dengue fever in brazil using species distribution model. *Trans. R. Soc. Trop. Med. Hyg.*, 108(2):99–104.

[Carlson et al., 2016] Carlson, C. J., Dougherty, E. R., and Getz, W. (2016). An ecological assessment of the pandemic threat of zika virus. *PLoS Negl. Trop. Dis.*, 10(8):e0004968.

[Carrington et al., 2013] Carrington, L. B., Armijos, M. V., Lambrechts, L., Barker, C. M., and Scott, T. W. (2013). Effects of fluctuating daily temperatures at critical thermal extremes on aedes aegypti life-history traits. *PLoS One*, 8(3):e58824.

[Castro et al., 2019] Castro, M. C., Baeza, A., Codeço, C. T., Cucunubá, Z. M., Dal'Asta, A. P., De Leo, G. A., Dobson, A. P., Carrasco-Escobar, G., Lana, R. M., Lowe, R., Monteiro, A. M. V., Pascual, M., and Santos-Vega, M. (2019). Development, environmental degradation, and disease spread in the brazilian amazon. *PLoS Biol.*, 17(11):e3000526.

[Cavender-Bares et al., 2016] Cavender-Bares, J., Meireles, J. E., Couture, J. J., Kaproth, M. A., Kingdon, C. C., Singh, A., Serbin, S. P., Center, A., Zuniga, E., Pilz, G., and Townsend, P. A. (2016). Associations of leaf spectra with genetic and phylogenetic variation in oaks: Prospects for remote detection of biodiversity. *Remote Sensing*, 8(3):221.

[Ceballos et al., 2017] Ceballos, G., Ehrlich, P. R., and Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad. Sci. U. S. A.*, 114(30):E6089–E6096.

[Chambers et al., 2009] Chambers, J. Q., Negrón-Juárez, R. I., Hurtt, G. C., Marra, D. M., and Higuchi, N. (2009). Lack of intermediate-scale disturbance data prevents robust extrapolation of plot-level tree mortality rates for old-growth tropical forests. *Ecol. Lett.*, 12(12).

[Chander et al., 2009] Chander, G., Markham, B. L., and Helder, D. L. (2009). Summary of current radiometric calibration coefficients for landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.*, 113(5):893–903.

[Chase and Knight, 2013] Chase, J. M. and Knight, T. M. (2013). Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. *Ecol. Lett.*, 16 Suppl 1:17–26.

[Chave et al., 2005] Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., and Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145(1):87–99.

[Chen and Kearney, 2015] Chen, Z. and Kearney, C. M. (2015). Nectar protein content and attractiveness to aedes aegypti and culex pipiens in plants with nectar/insect associations. *Acta Trop.*, 146:81–88.

[Ciais et al., 2014] Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., and Others (2014). Carbon and other biogeochemical cycles. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 465–570. Cambridge University Press.

[Clark et al., 2005] Clark, M. L., Roberts, D. A., and Clark, D. B. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens. Environ.*, 96(3):375–398.

[Cleveland et al., 2015] Cleveland, C. C., Taylor, P., Chadwick, K. D., Dahlin, K., Doughty, C. E., Malhi, Y., Smith, W. K., Sullivan, B. W., Wieder, W. R., and Townsend, A. R. (2015). A comparison of plot-based satellite and earth system model estimates of tropical forest net primary production. *Global Biogeochem. Cycles*, 29(5):626–644.

[Cohen et al., 2018] Cohen, W. B., Yang, Z., Healey, S. P., Kennedy, R. E., and Gorelick, N. (2018). A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sens. Environ.*, 205:131–140.

[Cohen et al., 2016] Cohen, W. B., Yang, Z., Stehman, S. V., Schroeder, T. A., Bell, D. M., Masek, J. G., Huang, C., and Meigs, G. W. (2016). Forest disturbance across the conterminous united states from 1985–2012: the emerging dominance of forest decline. *Forest Ecology and Management*, 360:242–252.

[Colgan et al., 2012] Colgan, M. S., Baldeck, C. A., Féret, J.-B., and Asner, G. P. (2012). Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data. *Remote Sensing*, 4(11):3462–3480.

[Collins et al., 2013] Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M. (2013). Chapter 12 - long-term climate change: Projections, commitments and irreversibility. In IPCC, editor, *Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5.* Cambridge University Press, Cambridge.

[Costa et al., 2010] Costa, E. A. P. d. A., Santos, E. M. d. M., Correia, J. C., and Albuquerque, C. M. R. d. (2010). Impact of small variations in temperature and humidity on the reproductive activity and survival of aedes aegypti (diptera, culicidae). *Rev. Bras. Entomol.*, 54(3):488–493.

[Culina et al., 2018] Culina, A., Baglioni, M., Crowther, T. W., Visser, M. E., Woutersen-Windhouwer, S., and Manghi, P. (2018). Navigating the unfolding open data landscape in ecology and evolution. *Nat Ecol Evol*, 2(3):420–426.

[Daily, 1999] Daily, G. (1999). Developing a scientific basis for managing earth's life support systems. *Conserv. Ecol.*, 3(2).

[Daily et al., 2009] Daily, G. C., Polasky, S., Goldstein, J., Kareiva, P. M., Mooney, H. A., Pejchar, L., Ricketts, T. H., Salzman, J., and Shallenberger, R. (2009). Ecosystem services in decision making: time to deliver. *Front. Ecol. Environ.*, 7(1):21–28.

[DeGroot and Fienberg, 1983] DeGroot, M. H. and Fienberg, S. E. (1983). The comparison and evaluation of forecasters. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 32(1/2):12–22.

[Delatte et al., 2008] Delatte, H., Dehecq, J. S., Thiria, J., Domerg, C., Paupy, C., and Fontenille, D. (2008). Geographic distribution and developmental sites of aedes albopictus (diptera: Culicidae) during a chikungunya epidemic event. *Vector Borne Zoonotic Dis.*, 8(1):25–34.

[der Walt et al., 2011] der Walt, S. v., Colbert, S. C., and Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science Engineering*, 13(2):22–30.

[Dietze et al., 2018] Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., and White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proc. Natl. Acad. Sci. U. S. A.*, 115(7):1424–1432.

[Diniz-Filho and Bini, 2019] Diniz-Filho, J. A. F. and Bini, L. M. (2019). Will life find a way out? evolutionary rescue and darwinian adaptation to climate change. *Perspectives in Ecology and Conservation*, 17(3):117–121.

[Dirzo et al., 2014] Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., and Collen, B. (2014). Defaunation in the anthropocene. *Science*, 345(6195):401–406.

[Donaldson et al., 2016] Donaldson, M. R., Burnett, N. J., Braun, D. C., Suski, C. D., Hinch, S. G., Cooke, S. J., and Kerr, J. T. (2016). Taxonomic bias and international biodiversity conservation research.

[Ehrlich and Mooney, 1983] Ehrlich, P. R. and Mooney, H. A. (1983). Extinction, substitution, and ecosystem services. *Bioscience*, 33(4):248–254.

[Elith et al., 2008] Elith, J., Leathwick, J. R., and Hastie, T. (2008). A working guide to boosted regression trees. *J. Anim. Ecol.*, 77(4):802–813.

[Elith et al., 2011] Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. *Diversity and Distributions*, 17(1):43–57.

[Ellis and Ramankutty, 2008] Ellis, E. C. and Ramankutty, N. (2008). Putting people in the map: anthropogenic biomes of the world. *Front. Ecol. Environ.*, 6(8):439–447.

[Engel et al., 2008] Engel, S., Pagiola, S., and Wunder, S. (2008). Designing payments for environmental services in theory and practice: An overview of the issues. *Ecol. Econ.*, 65(4):663–674.

[Essl et al., 2015] Essl, F., Dullinger, S., Rabitsch, W., Hulme, P. E., Pyšek, P., Wilson, J. R., and Richardson, D. M. (2015). Delayed biodiversity change: no time to waste. *Trends in Ecology & Evolution*, 30(7):375–378.

[Estes et al., 2018] Estes, L., Elsen, P. R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., Choi, J. J., and Ellis, E. C. (2018). The spatial and temporal domains of modern ecology. *Nature Ecology & Evolution*, 2(5):819–826.

[Faleiro et al., 2013] Faleiro, F. V., Machado, R. B., and Loyola, R. D. (2013). Defining spatial conservation priorities in the face of land-use and climate change. *Biological Conservation*, 158:248–257.

[Fan et al., 2015] Fan, H., Fu, X., Zhang, Z., and Wu, Q. (2015). Phenology-Based vegetation index differencing for mapping of rubber plantations using landsat OLI data. *Remote Sensing*, 7(5):6041–6058.

[FAO, 2016] FAO (2016). State of the world's forests 2016. forests and agriculture: land-use challenges and opportunities. Technical report, FAO.

[Fassnacht et al., 2016] Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., and Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.*, 186:64–87.

[Feilhauer et al., 2015] Feilhauer, H., Asner, G. P., and Martin, R. E. (2015). Multi-method ensemble selection of spectral bands related to leaf biochemistry. *Remote Sens. Environ.*, 164:57–65.

[Feng et al., 2008] Feng, J., Tang, H., Bai, L., Zhou, Q., and Chen, Z. (2008). New fast detection method of forest fire monitoring and application based on FY-1D/MVISR data. In *Computer And Computing Technologies In Agriculture, Volume I*, The International Federation for Information Processing, pages 613–628. Springer US.

[Fensholt et al., 2004] Fensholt, R., Sandholt, I., and Rasmussen, M. S. (2004). Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sens. Environ.*, 91(3):490–507.

[Feret and Asner, 2013] Feret, J. B. and Asner, G. P. (2013). Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Trans. Geosci. Remote Sens.*, 51(1):73–84.

[Féret and Asner, 2014] Féret, J.-B. and Asner, G. P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecol. Appl.*, 24(6):1289–1296.

[Fernández et al., 2015] Fernández, M., Navarro, L. M., Apaza-Quevedo, A., Gallegos, S. C., Marques, A., Zambrana-Torrelio, C., Wolf, F., Hamilton, H., Aguilar-Kirigin, A. J., Aguirre, L. F., Alvear, M., Aparicio, J., Apaza-Vargas, L., Arellano, G., Armijo, E., Ascarrunz, N., Barrera, S., Beck, S. G., Cabrera-Condarco, H., Campos-Villanueva, C., Cayola, L., Flores-Saldana, N. P., Fuentes, A. F., García-Lino, M. C., Gómez, M. I., Higueras, Y. S., Kessler, M., Ledezma, J. C., Limachi, J. M., López, R. P., Loza, M. I., Macía, M. J., Meneses, R. I., Miranda, T. B., Miranda-Calle, A. B., Molina-Rodriguez, R. F., Moraes R., M., Moya-Diaz, M. I., Ocampo, M., Perotto-Baldivieso, H. L., Plata, O., Reichle, S., Rivero, K., Seidel, R., Soria, L., Terán, M. F., Toledo, M., Zenteno-Ruiz, F. S., and Pereira, H. M. (2015). Challenges and opportunities for the bolivian biodiversity observation network. *Biodiversity*, 16(2-3):86–98.

[Ferraro and Pattanayak, 2006] Ferraro, P. J. and Pattanayak, S. K. (2006). Money for nothing? a call for empirical evaluation of biodiversity conservation investments. *PLoS Biol.*, 4(4):e105.

[Fick and Hijmans, 2017] Fick, S. E. and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.*, 37(12):4302–4315.

[Field, 1991] Field, C. B. (1991). Ecological scaling of carbon gain to stress and resource. In *Response of plants to multiple stresses*, pages 35–65. Academic Press San Diego.

[Field et al., 1995] Field, C. B., Randerson, J. T., and Malmström, C. M. (1995). Global net primary production: Combining ecology and remote sensing. *Remote Sens. Environ.*, 51(1):74–88.

[Fisher et al., 2008] Fisher, J. I., Hurtt, G. C., Thomas, R. Q., and Chambers, J. Q. (2008). Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.*, 11(6):554–563.

[Fisher, 1997] Fisher, P. (1997). The pixel: a snare and a delusion. *International Journal of Remote Sensing*, 18(3):679–685.

[Foley et al., 2005] Foley, J. A., Defries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K. (2005). Global consequences of land use. *Science*, 309(5734):570–574.

[Fourcade et al., 2018] Fourcade, Y., Besnard, A. G., and Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Glob. Ecol. Biogeogr.*, 27(2):245–256.

[Friedman, 2001] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Stat.*, 29(5):1189–1232.

[Fullen and Mitchell, 1994] Fullen, M. A. and Mitchell, D. J. (1994). Desertification and reclamation in North-Central china. *Ambio*, 23(2):131–135.

[Funk et al., 2017] Funk, J. L., Larson, J. E., Ames, G. M., Butterfield, B. J., Cavender-Bares, J., Firn, J., Laughlin, D. C., Sutton-Grier, A. E., Williams, L., and Wright, J. (2017). Revisiting the holy grail: using plant functional traits to understand ecological processes. *Biol. Rev. Camb. Philos. Soc.*, 92(2):1156–1173.

[Gairola et al., 2013] Gairola, S., Procheş, Ş., and Rocchini, D. (2013). High-resolution satellite remote sensing: a new frontier for biodiversity exploration in indian himalayan forests. *Int. J. Remote Sens.*, 34(6):2006–2022.

[Galeazzi et al., 2008] Galeazzi, C., Sacchetti, A., Cisbani, A., and Babini, G. (2008). The PRISMA program. In *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, volume 4, pages IV–105–IV − 108. ieeexplore.ieee.org.

[Gao et al., 2009] Gao, B.-C., Montes, M. J., Davis, C. O., and Goetz, A. F. H. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sens. Environ.*, 113:S17–S24.

[Garrigues et al., 2006] Garrigues, S., Allard, D., Baret, F., and Weiss, M. (2006). Influence of landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data. *Remote Sens. Environ.*, 105(4):286–298.

[Gaston et al., 2007] Gaston, K. J., Evans, K. L., and Lennon, J. J. (2007). The scaling of spatial turnover: pruning the thicket. *Scaling biodiversity*, pages 181–222.

[GBIF, 2018a] GBIF (2018a). Aedes aegypti species occurrence records, 2000-2017. Title of the publication associated with this dataset: GBIF.

[GBIF, 2018b] GBIF (2018b). Aedes albopictus species occurrence records, 2000-2017. Title of the publication associated with this dataset: GBIF.

[Geijzendorffer et al., 2016] Geijzendorffer, I. R., Regan, E. C., Pereira, H. M., Brotons, L., Brummitt, N., Gavish, Y., Haase, P., Martin, C. S., Mihoub, J.-B., Secades, C., Schmeller, D. S., Stoll, S., Wetzel, F. T., and Walters, M. (2016). Bridging the gap between biodiversity data and policy reporting needs: An essential biodiversity variables perspective. *J. Appl. Ecol.*, 53(5):1341–1350.

[Geist and Lambin, 2002] Geist, H. J. and Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation: Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. *Bioscience*, 52(2):143–150.

[GEO BON, 2017] GEO BON (2017). GEO BON strategy for development of essential biodiversity variables. Technical report, GEO BON.

[Gilbert et al., 2018] Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R. W., and Robinson, T. P. (2018). Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Sci Data*, 5:180227.

[Gillespie et al., 2008] Gillespie, T. W., Foody, G. M., Rocchini, D., Giorgi, A. P., and Saatchi, S. (2008). Measuring and modelling biodiversity from space. *Progress in Physical Geography: Earth and Environment*, 32(2):203–221.

[Godfray, 2007] Godfray, Jr, H. C. J. (2007). Linnaeus in the information age. *Nature*, 446(7133):259–260.

[Goetz et al., 1985] Goetz, A. F., Vane, G., Solomon, J. E., and Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *Science*, 228(4704):1147–1153.

[Gómez-Dantés and Willoquet, 2009] Gómez-Dantés, H. and Willoquet, J. R. (2009). Dengue in the americas: challenges for prevention and control. *Cad. Saude Publica*, 25 Suppl 1:S19–31.

[Gonzalez et al., 2016] Gonzalez, A., Cardinale, B. J., Allington, G. R. H., Byrnes, J., Arthur Endsley, K., Brown, D. G., Hooper, D. U., Isbell, F., O'Connor, M. I., and Loreau, M. (2016). Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. *Ecology*, 97(8):1949–1960.

[Gorelick et al., 2017] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.*, 202(Supplement C):18–27.

[Graesser et al., 2015] Graesser, J., Mitchell Aide, T., Ricardo Grau, H., and Ramankutty, N. (2015). Cropland/pastureland dynamics and the slowdown of deforestation in latin america. *Environ. Res. Lett.*, 10(3):034017.

[Gratz, 2004] Gratz, N. G. (2004). Critical review of the vector status of aedes albopictus. *Med. Vet. Entomol.*, 18(3):215–227.

[Grinnell, 1917] Grinnell, J. (1917). The Niche-Relationships of the california thrasher. *Auk*, 34(4):427–433.

[Gubler, 2002] Gubler, D. J. (2002). Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiol.*, 10(2):100–103.

[Guerry et al., 2015] Guerry, A. D., Polasky, S., Lubchenco, J., Chaplin-Kramer, R., Daily, G. C., Griffin, R., Ruckelshaus, M., Bateman, I. J., Duraiappah, A., Elmqvist, T., Feldman, M. W., Folke, C., Hoekstra, J., Kareiva, P. M., Keeler, B. L., Li, S., McKenzie, E., Ouyang, Z., Reyers, B., Ricketts, T. H., Rockström, J., Tallis, H., and Vira, B. (2015). Natural capital and ecosystem services informing decisions: From promise to practice. *Proc. Natl. Acad. Sci. U. S. A.*, 112(24):7348–7355.

[Guisan et al., 2007] Guisan, A., Graham, C. H., Elith, J., Huettmann, F., and the NCEAS Species Distribution Modelling Group (2007). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, 13(3):332–340.

[Guisan and Thuiller, 2005] Guisan, A. and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.*, 8(9):993–1009.

[Gutiérrez, 2015] Gutiérrez, L. A. (2015). PAHO/WHO data - dengue cases — PAHO/WHO. `http://www.paho.org/data/index.php/en/mnu-topics/indicadores-dengue-en/dengue-nacional-en/252-dengue-pais-ano-en.html`. Accessed: 2019-5-6.

[Hall and Llinas, 1997] Hall, D. L. and Llinas, J. (1997). An introduction to multisensor data fusion. *Proc. IEEE*, 85(1):6–23.

[Hallé et al., 1978] Hallé, F., Oldeman, R. A. A., and Tomlinson, P. B. (1978). *Tropical Trees and Forests: An Architectural Analysis*. Springer, Berlin, Heidelberg.

[Hamrick, 2016] Hamrick, K. (2016). State of private investment in conservation 2016: A landscape assessment of an emerging market. *Forest Trends*.

[Hansen et al., 2013] 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., and Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853.

[Hastie and Tibshirani, 2004] Hastie, T. and Tibshirani, R. (2004). Generalized additive models. In *Encyclopedia of Statistical Sciences*. John Wiley & Sons, Inc.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., and Friedman, J. (2009). Unsupervised learning. In Hastie, T., Tibshirani, R., and Friedman, J., editors, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, pages 485–585. Springer New York, New York, NY.

[Hawkins, 2012] Hawkins, B. A. (2012). Eight (and a half) deadly sins of spatial analysis: Spatial analysis. *J. Biogeogr.*, 39(1):1–9.

[Henderson and Lewis, 2008] Henderson, F. M. and Lewis, A. J. (2008). Radar detection of wetland ecosystems: a review. *Int. J. Remote Sens.*, 29(20):5809–5835.

[Hesketh and Sánchez-Azofeifa, 2012] Hesketh, M. and Sánchez-Azofeifa, G. A. (2012). The effect of seasonal spectral variation on species classification in the panamanian tropical forest. *Remote Sens. Environ.*, 118:73–82.

[Hilker et al., 2009] Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., Gao, F., and White, J. C. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on landsat and MODIS. *Remote Sens. Environ.*, 113(8):1613–1627.

[Hopp and Foley, 2001] Hopp, M. J. and Foley, J. A. (2001). Global-Scale relationships between climate and the dengue fever vector, aedes aegypti. *Clim. Change*, 48(2):441–463.

[Hortal et al., 2015] Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., and Ladle, R. J. (2015). Seven shortfalls that beset Large-Scale knowledge of biodiversity. *Annu. Rev. Ecol. Evol. Syst.*, 46(1):523–549.

[Hou et al., 2013] Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2013). The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.*, 95(5):701–722.

[Hubbell, 2001] Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press.

[Huffman et al., 2007] Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-Global, multiyear, Combined-Sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, 8(1):38–55.

[Hungate et al., 2017] Hungate, B. A., Barbier, E. B., Ando, A. W., Marks, S. P., Reich, P. B., van Gestel, N., Tilman, D., Knops, J. M. H., Hooper, D. U., Butterfield, B. J., and Cardinale, B. J. (2017). The economic value of grassland species for carbon storage. *Sci Adv*, 3(4):e1601880.

[Hunter, 2007] Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science Engineering*, 9(3):90–95.

[Hurlbert and Jetz, 2007] Hurlbert, A. H. and Jetz, W. (2007). Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proc. Natl. Acad. Sci. U. S. A.*, 104(33):13384–13389.

[Hutchinson, 1953] Hutchinson, G. E. (1953). The concept of pattern in ecology. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 105:1–12.

[Immitzer et al., 2012] Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, 4(9):2661–2693.

[IPBES, 2019] IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Technical Report 1, IPBES Secretariat.

[Irons et al., 2012] Irons, J. R., Dwyer, J. L., and Barsi, J. A. (2012). The next landsat satellite: The landsat data continuity mission. *Remote Sens. Environ.*, 122(Supplement C):11–21.

[Iyengar and Massey, 2018] Iyengar, S. and Massey, D. S. (2018). Scientific communication in a post-truth society. *Proc. Natl. Acad. Sci. U. S. A.*

[Jacquemoud et al., 2009] Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., and Ustin, S. L. (2009). Prospect+ sail models: A review of use for vegetation characterization. *Remote sensing of environment*, 113:S56–S66.

[Jansen and Beebe, 2010] Jansen, C. C. and Beebe, N. W. (2010). The dengue vector aedes aegypti: what comes next. *Microbes Infect.*, 12(4):272–279.

[Jensen and Lulla, 1987] Jensen, J. R. and Lulla, K. (1987). Introductory digital image processing: A remote sensing perspective. *Geocarto Int.*, 2(1):65–65.

[Jetz et al., 2016] 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., and Ustin, S. L. (2016). Monitoring plant functional diversity from space. *Nat Plants*, 2(3):16024.

[Jetz et al., 2012] Jetz, W., McPherson, J. M., and Guralnick, R. P. (2012). Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol. Evol.*, 27(3):151–159.

[Jia and Richards, 1999] Jia, X. and Richards, J. A. (1999). Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Trans. Geosci. Remote Sens.*, 37(1):538–542.

[Jones et al., 2018] Jones, K. R., Venter, O., Fuller, R. A., Allan, J. R., Maxwell, S. L., Negret, P. J., and Watson, J. E. M. (2018). One-third of global protected land is under intense human pressure. *Science*, 360(6390):788–791.

[Jordan, 1969] Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 50(4):663–666.

[Justice et al., 1998] Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., Strahler, A., Lucht, W., Myneni, R. B., Knyazikhin, Y., Running, S. W., Nemani, R. R., Wan, Z., Huete, A. R., van Leeuwen, W., Wolfe, R. E., Giglio, L., Muller, J., Lewis, P., and Barnsley, M. J. (1998). The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.*, 36(4):1228–1249.

[Kamgang et al., 2012] Kamgang, B., Nchoutpouen, E., Simard, F., and Paupy, C. (2012). Notes on the blood-feeding behavior of aedes albopictus (diptera: Culicidae) in cameroon. *Parasit. Vectors*, 5:57.

[Kampe et al., 2010] Kampe, T. U., Johnson, B. R., Kuester, M. A., and Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *JARS*, 4(1):043510.

[Kattge et al., 2011] Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van BODEGOM, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., Atkin, O., Bahn, M., Baker, T. R., Baldocchi, D., Bekker, R., Blanco, C. C., Blonder, B., Bond, W. J., Bradstock, R., Bunker, D. E., Casanoves, F., Cavender-Bares, J., Chambers, J. Q., Chapin, Iii, F. S., Chave, J., Coomes, D., Cornwell, W. K., Craine, J. M., Dobrin, B. H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte, M., Fagan, W. F., Fang, J., Fernández-Méndez, F., Fidelis, A., Finegan, B., Flores, O., Ford, H., Frank, D., Freschet, G. T., Fyllas, N. M., Gallagher, R. V., Green, W. A., Gutierrez, A. G., Hickler, T., Higgins, S. I., Hodgson, J. G., Jalili, A., Jansen, S., Joly, C. A., Kerkhoff, A. J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J. M. H., Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T. D., Leishman, M., Lens, F., Lenz, T., Lewis, S. L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M. D., Manning, P., Massad, T., Medlyn, B. E., Messier, J., Moles, A. T., Müller, S. C., Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn, J., Onipchenko, V. G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W. A., Patiño, S., Paula, S., Pausas, J. G., Peñuelas, J., Phillips, O. L., Pillar, V., Poorter, H., Poorter, L., Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L., Salgado-Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E., Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M., Weiher, E., White, M., White, S., Wright, S. J., Yguel, B., Zaehle, S., Zanne, A. E., and Wirth, C. (2011). TRY - a global database of plant traits. *Glob. Chang. Biol.*, 17(9):2905–2935.

[Keil et al., 2012] Keil, P., Schweiger, O., Kühn, I., Kunin, W. E., Kuussaari, M., Settele, J., Henle, K., Brotons, L., Pe'er, G., Lengyel, S., Moustakas, A., Steinicke, H., and Storch, D. (2012). Patterns of beta diversity in europe: the role of climate, land cover and distance across scales: Multiscale patterns of beta diversity in europe. *J. Biogeogr.*, 39(8):1473–1486.

[Keil et al., 2015] Keil, P., Storch, D., and Jetz, W. (2015). On the decline of biodiversity due to area loss. *Nat. Commun.*, 6:8837.

[Keller et al., 2008] Keller, M., Schimel, D. S., Hargrove, W. W., and Hoffman, F. M. (2008). A continental strategy for the national ecological observatory network. *Front. Ecol. Environ.*, 6(5):282–284.

[Khare et al., 2018] Khare, S., Latifi, H., and Ghosh, S. K. (2018). Multi-scale assessment of invasive plant species diversity using pléiades 1a, RapidEye and landsat-8 data. *Geocarto Int.*, 33(7):681–698.

[Kichenin et al., 2013] Kichenin, E., Wardle, D. A., Peltzer, D. A., Morse, C. W., and Freschet, G. T. (2013). Contrasting effects of plant inter-and intraspecific variation on community-level trait measures along an environmental gradient. *Funct. Ecol.*, 27(5):1254–1261.

[Knudsen, 1995] Knudsen, A. B. (1995). Global distribution and continuing spread of aedes albopictus. *Parassitologia*, 37(2-3):91–97.

[Kogan et al., 2015] Kogan, F., Goldberg, M., Schott, T., and Guo, W. (2015). Suomi NPP/VIIRS: improving drought watch, crop loss prediction, and food security. *Int. J. Remote Sens.*, 36(21):5373–5383.

[Kokaly et al., 2009] Kokaly, R. F., Asner, G. P., Ollinger, S. V., Martin, M. E., and Wessman, C. A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sens. Environ.*, 113(Supplement 1):S78–S91.

[Konings et al., 2019] Konings, A. G., Rao, K., and Steele-Dunne, S. C. (2019). Macro to micro: microwave remote sensing of plant water content for physiology and ecology. *New Phytol.*, 223(3):1166–1172.

[Korzukhin et al., 1996] Korzukhin, M. D., Ter-Mikaelian, M. T., and Wagner, R. G. (1996). Process versus empirical models: which approach for forest ecosystem management? *Canadian Journal of Forest Research*, 26(5):879–887.

[Kraemer et al., 2015a] Kraemer, M. U. G., Sinka, M. E., Duda, K. A., Mylne, A., Shearer, F. M., Brady, O. J., Messina, J. P., Barker, C. M., Moore, C. G., Carvalho, R. G., Coelho, G. E., Van Bortel, W., Hendrickx, G., Schaffner, F., Wint, G. R. W., Elyazar, I. R. F., Teng, H.-J., and Hay, S. I. (2015a). Data from: The global compendium of aedes aegypti and ae. albopictus occurrence.

[Kraemer et al., 2015b] Kraemer, M. U. G., Sinka, M. E., Duda, K. A., Mylne, A. Q. N., Shearer, F. M., Barker, C. M., Moore, C. G., Carvalho, R. G., Coelho, G. E., Van Bortel, W., Hendrickx, G., Schaffner, F., Elyazar, I. R. F., Teng, H.-J., Brady, O. J., Messina, J. P., Pigott, D. M., Scott, T. W., Smith, D. L., Wint, G. R. W., Golding, N., and Hay, S. I. (2015b). The global distribution of the arbovirus vectors aedes aegypti and ae. albopictus. *Elife*, 4:e08347.

[Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc.

[Kröger, 2020] Kröger, M. (2020). Deforestation, cattle capitalism and neodevelopmentalism in the chico mendes extractive reserve, brazil. *J. Peasant Stud.*, 47(3):464–482.

[Kruse et al., 2011] Kruse, F. A., Taranik, J. V., Coolbaugh, M., Michaels, J., Littlefield, E. F., Calvin, W. M., and Martini, B. A. (2011). Effect of reduced spatial resolution on mineral mapping using imaging Spectrometry—Examples using hyperspectral infrared imager (HyspIRI)-Simulated data. *Remote Sensing*, 3(8):1584–1602.

[Kurtzer et al., 2017] Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS One*, 12(5):e0177459.

[Landau and van Leeuwen, 2012] Landau, K. I. and van Leeuwen, W. J. D. (2012). Fine scale spatial urban land cover factors associated with adult mosquito abundance and risk in tucson, arizona. *J. Vector Ecol.*, 37(2):407–418.

[Larigauderie and Mooney, 2010] Larigauderie, A. and Mooney, H. A. (2010). The intergovernmental science-policy platform on biodiversity and ecosystem services: moving a step closer to an IPCC-like mechanism for biodiversity. *Current Opinion in Environmental Sustainability*, 2(1):9–14.

[Laurance et al., 2012] Laurance, W. F., Useche, D. C., 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., Kress, W. J., 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., Urbina-Cardona, J. N., Vasudevan, K., Wright, S. J., Arias-G, J. C., 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., Beisiegel, B. d. M., da Silva, R. d. N. P., Di Fiore, A., Diesmos, A., Dirzo, R., Doran-Sheehy, D., Eaton, M., Emmons, L., Estrada, A., Ewango, C., Fedigan, L., Feer, F., Fruth, B., Willis, J. G., Goodale, U., Goodman, S., Guix, J. C., Guthiga, P., Haber, W., Hamer, K., Herbinger,

I., Hill, J., Huang, Z., Sun, I. F., 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.-B., Klop, E., Knott, C., Koné, I., Kudavidanage, E., Ribeiro, J. L. d. S., Lattke, J., Laval, R., Lawton, R., Leal, M., Leighton, M., Lentino, M., Leonel, C., Lindsell, J., Ling-Ling, L., Linsenmair, K. E., Losos, E., Lugo, A., Lwanga, J., Mack, A. L., Martins, M., McGraw, W. S., McNab, R., Montag, L., Thompson, J. M., 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., Rivera, J. V., Venkataraman, M., Venn, L., Verea, C., de Castilho, C. V., 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., and Zamzani, F. (2012). Averting biodiversity collapse in tropical forest protected areas. *Nature*, 489(7415):290–294.

[Lausch et al., 2016] 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., and Cord, A. F. (2016). Linking earth observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. *Ecol. Indic.*, 70(Supplement C):317–339.

[Lechner, 2010] Lechner, A. (2010). *Scale in remote sensing and its impact on landscape ecology.* PhD thesis, RMIT University.

[Lee et al., 1994] Lee, J.-S., Jurkevich, L., Dewaele, P., Wambacq, P., and Oosterlinck, A. (1994). Speckle filtering of synthetic aperture radar images: A review. *Remote sensing reviews*, 8(4):313–340.

[Lefsky et al., 2005] Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., Hunter, M. O., and de Oliveira, Jr., R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys. Res. Lett.*, 32(22):L22S02.

[Legendre et al., 2005] Legendre, P., Borcard, D., and Peres-Neto, P. R. (2005). Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecol. Monogr.*, 75(4):435–450.

[Lepine et al., 2016] Lepine, L. C., Ollinger, S. V., Ouimette, A. P., and Martin, M. E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sens. Environ.*, 173:174–186.

[Leutner et al., 2012] Leutner, B. F., Reineking, B., Müller, J., Bachmann, M., Beierkuhnlein, C., Dech, S., and Wegmann, M. (2012). Modelling forest α-Diversity and floristic composition — on the added value of LiDAR plus hyperspectral remote sensing. *Remote Sensing*, 4(9):2818–2845.

[Levin, 1992] Levin, S. A. (1992). The problem of pattern and scale in ecology: The robert h. MacArthur award lecture. *Ecology*, 73(6):1943–1967.

[Li et al., 2014] Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., Zhou, Y., Yao, L., Yan, G., and Chen, X.-G. (2014). Urbanization increases aedes albopictus larval habitats and accelerates mosquito development and survivorship. *PLoS Negl. Trop. Dis.*, 8(11):e3301.

[Li et al., 2013] Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.*, 131:14–37.

[Liang et al., 2016] Liang, J., Crowther, T. W., Picard, N., Wiser, S., Zhou, M., Alberti, G., Schulze, E.-D., McGuire, A. D., Bozzato, F., Pretzsch, H., de Miguel, S., Paquette, A., Hérault, B., Scherer-Lorenzen, M., Barrett, C. B., Glick, H. B., Hengeveld, G. M., Nabuurs, G.-J., Pfautsch, S., Viana, H., Vibrans, A. C., Ammer, C., Schall, P., Verbyla, D., Tchebakova, N., Fischer, M., Watson, J. V., Chen, H. Y. H., Lei, X., Schelhaas, M.-J., Lu, H., Gianelle, D., Parfenova, E. I., Salas, C., Lee, E., Lee, B., Kim, H. S., Bruelheide, H., Coomes, D. A., Piotto, D., Sunderland, T., Schmid, B., Gourlet-Fleury, S., Sonké, B., Tavani, R., Zhu, J., Brandl, S., Vayreda, J., Kitahara, F., Searle, E. B., Neldner, V. J., Ngugi, M. R., Baraloto, C., Frizzera, L., Bałazy, R., Oleksyn, J., Zawiła-Niedźwiecki, T., Bouriaud, O., Bussotti, F., Finér, L., Jaroszewicz, B., Jucker, T., Valladares, F., Jagodzinski, A. M., Peri, P. L., Gonmadje, C., Marthy, W., O'Brien, T., Martin, E. H., Marshall, A. R., Rovero, F., Bitariho, R., Niklaus, P. A., Alvarez-Loayza, P., Chamuya, N., Valencia, R., Mortier, F., Wortel, V., Engone-Obiang, N. L., Ferreira, L. V., Odeke, D. E., Vasquez, R. M., Lewis, S. L., and Reich, P. B. (2016). Positive biodiversity-productivity relationship predominant in global forests. *Science*, 354(6309).

[Liu et al., 2020] Liu, C., Wolter, C., Xian, W., and Jeschke, J. M. (2020). Species distribution models have limited spatial transferability for invasive species. *Ecol. Lett.*

[Liu-Helmersson et al., 2014] Liu-Helmersson, J., Stenlund, H., Wilder-Smith, A., and Rocklöv, J. (2014). Vectorial capacity of aedes aegypti: effects of temperature and implications for global dengue epidemic potential. *PLoS One*, 9(3):e89783.

[Locatelli et al., 2011] Locatelli, B., Evans, V., Wardell, A., Andrade, A., and Vignola, R. (2011). Forests and climate change in latin america: Linking adaptation and mitigation. *For. Trees Livelihoods*, 2(1):431–450.

[Mac Nally et al., 2004] Mac Nally, R., Fleishman, E., Bulluck, L. P., and Betrus, C. J. (2004). Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies: Spatial scale and beta diversity in birds and butterflies. *J. Biogeogr.*, 31(6):917–929.

[MacArthur and Wilson, 1967] MacArthur, R. H. and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press.

[Maillard et al., 2008] Maillard, P., Alencar-Silva, T., and Clausi, D. A. (2008). An evaluation of radarsat-1 and ASTER data for mapping veredas (palm swamps). *Sensors*, 8(9):6055–6076.

[Malenovský et al., 2012] Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R., and Berger, M. (2012). Sentinels for science: Potential of sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sens. Environ.*, 120(Supplement C):91–101.

[Marconi et al., 2018] Marconi, S., Graves, S. J., Gong, D., Nia, M. S., Le Bras, M., Dorr, B. J., Fontana, P., Gearhart, J., Greenberg, C., Harris, D. J., Kumar, S. A., Nishant, A., Prarabdh, J., Rege, S. U., Bohlman, S. A., White, E. P., and Wang, D. Z. (2018). A data science challenge for converting airborne remote sensing data into ecological information. Technical Report e26966v1, PeerJ Preprints.

[Marengo et al., 2011] Marengo, J. A., Tomasella, J., Alves, L. M., Soares, W. R., and Rodriguez, D. A. (2011). The drought of 2010 in the context of historical droughts in the amazon region: DROUGHT AMAZON 2010. *Geophys. Res. Lett.*, 38(12):L12703.

[Marín Rodríguez et al., 2014] Marín Rodríguez, R., Calderón-Arguedas, O., Díaz Ríos, M., Duarte Solano, G., Valle Arguedas, J. J., and Troyo Rodríguez, A. (2014). First finding of aedes albopictus skuse in the the greater metropolitan area of costa rica. *Revista Costarricense de Salud Pública*, 23(1):01–04.

[Marinotti et al., 1990] Marinotti, O., James, A. A., and Ribeiro, J. C. (1990). Diet and salivation in female aedes aegypti mosquitoes. *J. Insect Physiol.*, 36(8):545–548.

[Martin et al., 2008] Martin, M. E., Plourde, L. C., Ollinger, S. V., Smith, M.-L., and McNeil, B. E. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sens. Environ.*, 112(9):3511–3519.

[Martin et al., 2018] Martin, R. E., Chadwick, K. D., Brodrick, P. G., Carranza-Jimenez, L., Vaughn, N. R., and Asner, G. P. (2018). An approach for foliar trait retrieval from airborne imaging spectroscopy of tropical forests. *Remote Sensing*, 10(2):199.

[Martinez, 1996] Martinez, N. D. (1996). Defining and measuring functional aspects of biodiversity. *Biodiversity. A biology of numbers of difference*, pages 114–148.

[Martinez-Ibarra et al., 1997] Martinez-Ibarra, J. A., Rodriguez, M. H., Arredondo-Jimenez, J. I., and Yuval, B. (1997). Influence of plant abundance on nectar feeding by aedes aegypti (diptera: Culicidae) in southern mexico. *J. Med. Entomol.*, 34(6):589–593.

[Marvin et al., 2014] Marvin, D. C., Asner, G. P., Knapp, D. E., Anderson, C. B., Martin, R. E., Sinca, F., and Tupayachi, R. (2014). Amazonian landscapes and the bias in field studies of forest structure and biomass. *Proc. Natl. Acad. Sci. U. S. A.*, 111(48):E5224–32.

[Mascaro et al., 2014a] Mascaro, J., Asner, G. P., Davies, S., Dehgan, A., and Saatchi, S. (2014a). These are the days of lasers in the jungle. *Carbon Balance Manag.*, 9(1):7.

[Mascaro et al., 2014b] Mascaro, J., Asner, G. P., Knapp, D. E., Kennedy-Bowdoin, T., Martin, R. E., Anderson, C., Higgins, M., and Chadwick, K. D. (2014b). A tale of two "forests": random forest machine learning AIDS tropical forest carbon mapping. *PLoS One*, 9(1):e85993.

[Matsunaga et al., 2011] Matsunaga, T., Yamamoto, S., Kashimura, O., Tachikawa, T., Ogawa, K., Iwasaki, A., Tsuchida, S., and Ohgi, N. (2011). Operation plan study for japanese future hyperspectral mission: HISUI. In *Proc. 34th International Symposium on Remote Sensing of Environment. Tucson: ISRSE.* Citeseer.

[Mayor et al., 2009] Mayor, S. J., Schneider, D. C., Schaefer, J. A., and Mahoney, S. P. (2009). Habitat selection at multiple scales. *Ecoscience*, 16(2):238–247.

[McGarigal et al., 2016] McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C., and Cushman, S. A. (2016). Multi-scale habitat selection modeling: a review and outlook. *Landscape ecology*, 31(6):1161–1175.

[McGill, 2010] McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecol. Lett.*, 13(5):627–642.

[McGill et al., 2015] McGill, B. J., Dornelas, M., Gotelli, N. J., and Magurran, A. E. (2015). Fifteen forms of biodiversity trend in the anthropocene. *Trends in ecology & evolution*, 30(2):104–113.

[McKinney and Others, 2010] McKinney, W. and Others (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56.

[McManus et al., 2016] McManus, K. M., Asner, G. P., Martin, R. E., Dexter, K. G., Kress, W. J., and Field, C. B. (2016). Phylogenetic structure of foliar spectral traits in tropical forest canopies. *Remote Sensing*, 8(3):196.

[Medley, 2010] Medley, K. A. (2010). Niche shifts during the global invasion of the asian tiger mosquito, aedes albopictus skuse (culicidae), revealed by reciprocal distribution models : Niche shifts and global invasion. *Glob. Ecol. Biogeogr.*, 19(1):122–133.

[Mendenhall et al., 2014] Mendenhall, C. D., Karp, D. S., Meyer, C. F. J., Hadly, E. A., and Daily, G. C. (2014). Predicting biodiversity change and averting collapse in agricultural landscapes. *Nature*, 509(7499):213–217.

[Mendenhall et al., 2011] Mendenhall, C. D., Sekercioglu, C. H., Brenes, F. O., Ehrlich, P. R., and Daily, G. C. (2011). Predictive model for sustaining biodiversity in tropical countryside. *Proc. Natl. Acad. Sci. U. S. A.*, 108(39):16313–16316.

[Mendenhall et al., 2016] Mendenhall, C. D., Shields-Estrada, A., Krishnaswami, A. J., and Daily, G. C. (2016). Quantifying and sustaining biodiversity in tropical agricultural landscapes. *Proc. Natl. Acad. Sci. U. S. A.*, 113(51):14544–14551.

[Merow et al., 2013] Merow, C., Smith, M. J., and Silander, Jr, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, 36(10):1058–1069.

[Metzger et al., 2013] Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A., and Zomer, R. (2013). A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. *Glob. Ecol. Biogeogr.*, 22(5):630–638.

[Mills et al., 2013] Mills, S., Weiss, S., and Liang, C. (2013). VIIRS day/night band (DNB) stray light characterization and correction. In *Earth Observing Systems XVIII*, volume 8866, page 88661P. International Society for Optics and Photonics.

[Ministerio de Salud, Perú, 2020] Ministerio de Salud, Perú (2020). MINSA declarará en emergencia sanitaria a regiones loreto, madre de dios y san martín para combatir brote de dengue. `https://www.gob.pe/institucion/minsa/noticias/` `81315-minsa-declarara-en-emergencia-sanitaria-a-regiones-loreto-madre-de-dios-y-san-martin-para-combatir` Accessed: 2020-3-17.

[Moore and Mitchell, 1997] Moore, C. G. and Mitchell, C. J. (1997). Aedes albopictus in the united states: ten-year presence and public health implications. *Emerg. Infect. Dis.*, 3(3):329–334.

[Mordecai et al., 2019] Mordecai, E. A., Caldwell, J. M., Grossman, M. K., Lippi, C. A., Johnson, L. R., Neira, M., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., Sippy, R., Stewart Ibarra, A. M., Thomas, M. B., and Villena, O. (2019). Thermal biology of mosquito-borne disease. *Ecol. Lett.*

[Mordecai et al., 2017] Mordecai, E. A., Cohen, J. M., Evans, M. V., Gudapati, P., Johnson, L. R., Lippi, C. A., Miazgowicz, K., Murdock, C. C., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., Stewart Ibarra, A., Thomas, M. B., and Weikel, D. P. (2017). Detecting the impact of temperature on transmission of zika, dengue, and chikungunya using mechanistic models. *PLoS Negl. Trop. Dis.*, 11(4):e0005568.

[Mordecai et al., 2013] Mordecai, E. A., Paaijmans, K. P., Johnson, L. R., Balzer, C., Ben-Horin, T., de Moor, E., McNally, A., Pawar, S., Ryan, S. J., Smith, T. C., and Lafferty, K. D. (2013). Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecol. Lett.*, 16(1):22–30.

[Morrison et al., 2012] Morrison, M. L., Marcot, B., and Mannan, W. (2012). *Wildlife-habitat relationships: concepts and applications*. Island Press.

[Mweya et al., 2016] Mweya, C. N., Kimera, S. I., Stanley, G., Misinzo, G., and Mboera, L. E. G. (2016). Climate change influences potential distribution of infected aedes aegypti Co-Occurrence with dengue epidemics risk areas in tanzania. *PLoS One*, 11(9):e0162649.

[Nagendra et al., 2013] Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., and Mairota, P. (2013). Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.*, 33:45–59.

[Nagendra et al., 2010] Nagendra, H., Rocchini, D., Ghate, R., Sharma, B., and Pareeth, S. (2010). Assessing plant diversity in a dry tropical forest: Comparing the utility of landsat and ikonos satellite images. *Remote Sensing*, 2(2):478–496.

[Naidoo et al., 2016] Naidoo, L., Mathieu, R., Main, R., Wessels, K., and Asner, G. P. (2016). L-band synthetic aperture radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs. *Int. J. Appl. Earth Obs. Geoinf.*, 52:54–64.

[Nekola and White, 1999] Nekola, J. C. and White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *J. Biogeogr.*, 26(4):867–878.

[Nemani et al., 2011] Nemani, R., Votava, P., Michaelis, A., Melton, F., and Milesi, C. (2011). Collaborative supercomputing for global change science. *Eos Trans. AGU*, 92(13):109.

[Nepstad et al., 2014a] Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M. C., McGrath-Horn, M., Carvalho, O., and Hess, L. (2014a). Slowing amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–1123.

[Nepstad et al., 2014b] Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M. C., McGrath-Horn, M., Carvalho, O., and Hess, L. (2014b). Slowing amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–1123.

[Niculescu-Mizil and Caruana, 2005] Niculescu-Mizil, A. and Caruana, R. (2005). Predicting good probabilities with supervised learning. In *Proceedings of the 22Nd International Conference on Machine Learning*, ICML '05, pages 625–632, New York, NY, USA. ACM.

[Nobre et al., 2016] Nobre, C. A., Sampaio, G., Borma, L. S., Castilla-Rubio, J. C., Silva, J. S., and Cardoso, M. (2016). Land-use and climate change risks in the amazon and the need of a novel sustainable development paradigm. *Proc. Natl. Acad. Sci. U. S. A.*, 113(39):10759–10768.

[Norris, 2004] Norris, D. E. (2004). Mosquito-borne diseases as a consequence of land use change. *Ecohealth*, 1(1):19–24.

[O'Connor et al., 2015] O'Connor, B., Secades, C., Penner, J., Sonnenschein, R., Skidmore, A., Burgess, N. D., and Hutton, J. M. (2015). Earth observation as a tool for tracking progress towards the aichi biodiversity targets. *Remote Sens Ecol Conserv*, 1(1):19–28.

[Oksanen, 1996] Oksanen, J. (1996). Is the humped relationship between species richness and biomass an artefact due to plot size? *J. Ecol.*, 84(2):293–295.

[Oliphant, 2007] Oliphant, T. E. (2007). Python for scientific computing. *Computing in Science Engineering*, 9(3):10–20.

[Ollinger, 2011] Ollinger, S. V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.*, 189(2):375–394.

[Ortega-Morales and Rodríguez, 2016] Ortega-Morales, A. I. and Rodríguez, Q. K. S. (2016). First record of aedes albopictus (diptera: Culicidae) in san luis potosi, mexico. *J. Vector Ecol.*, 41(2):314–315.

[Osnas et al., 2013] Osnas, J. L. D., Lichstein, J. W., Reich, P. B., and Pacala, S. W. (2013). Global leaf trait relationships: mass, area, and the leaf economics spectrum. *Science*, 340(6133):741–744.

[Ouyang et al., 2016] Ouyang, Z., Zheng, H., Xiao, Y., Polasky, S., Liu, J., Xu, W., Wang, Q., Zhang, L., Xiao, Y., Rao, E., Jiang, L., Lu, F., Wang, X., Yang, G., Gong, S., Wu, B., Zeng, Y., Yang, W., and Daily, G. C. (2016). Improvements in ecosystem services from investments in natural capital. *Science*, 352(6292):1455–1459.

[Paaijmans et al., 2013] Paaijmans, K. P., Heinig, R. L., Seliga, R. A., Blanford, J. I., Blanford, S., Murdock, C. C., and Thomas, M. B. (2013). Temperature variation makes ectotherms more sensitive to climate change. *Glob. Chang. Biol.*, 19(8):2373–2380.

[Padmanabha et al., 2012] Padmanabha, H., Durham, D., Correa, F., Diuk-Wasser, M., and Galvani, A. (2012). The interactive roles of aedes aegypti super-production and human density in dengue transmission. *PLoS Negl. Trop. Dis.*, 6(8):e1799.

[Pal, 2005] Pal, M. (2005). Random forest classifier for remote sensing classification. *Int. J. Remote Sens.*, 26(1):217–222.

[Palmer and White, 1994] Palmer, M. W. and White, P. S. (1994). Scale dependence and the Species-Area relationship. *Am. Nat.*, 144(5):717–740.

[Pan American Health Organization / World Health Organization, 2020] Pan American Health Organization / World Health Organization (2020). Epidemiological update: Dengue. Technical Report 1, PAHO / WHO.

[Papeş et al., 2010] Papeş, M., Tupayachi, R., Martínez, P., Peterson, A. T., and Powell, G. V. N. (2010). Using hyperspectral satellite imagery for regional inventories: a test with tropical emergent trees in the amazon basin. *J. Veg. Sci.*, 21(2):342–354.

[Paupy et al., 2009] Paupy, C., Delatte, H., Bagny, L., Corbel, V., and Fontenille, D. (2009). Aedes albopictus, an arbovirus vector: from the darkness to the light. *Microbes Infect.*, 11(14-15):1177–1185.

[Pech-May et al., 2016] Pech-May, A., Moo-Llanes, D. A., Puerto-Avila, M. B., Casas, M., Danis-Lozano, R., Ponce, G., Tun-Ku, E., Pinto-Castillo, J. F., Villegas, A., Ibáñez-Piñon, C. R., González, C., and Ramsey, J. M. (2016). Population genetics and ecological niche of invasive aedes albopictus in mexico. *Acta Trop.*, 157:30–41.

[Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, 12(Oct):2825–2830.

[Peluso and Lund, 2011] Peluso, N. L. and Lund, C. (2011). New frontiers of land control: Introduction. *J. Peasant Stud.*, 38(4):667–681.

[Pereira et al., 2013] 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., and Wegmann, M. (2013). Ecology. essential biodiversity variables. *Science*, 339(6117):277–278.

[Pereira et al., 2012] Pereira, H. M., Navarro, L. M., and Martins, I. S. (2012). Global biodiversity change: The bad, the good, and the unknown. *Annual Review of*.

[Peterson et al., 2005] Peterson, A. T., Martínez-Campos, C., Nakazawa, Y., and Martínez-Meyer, E. (2005). Time-specific ecological niche modeling predicts spatial dynamics of vector insects and human dengue cases. *Trans. R. Soc. Trop. Med. Hyg.*, 99(9):647–655.

[Pettorelli et al., 2014a] Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., and Turner, W. (2014a). Satellite remote sensing for applied ecologists: opportunities and challenges. *J. Appl. Ecol.*, 51(4):839–848.

[Pettorelli et al., 2016] Pettorelli, N., Owen, H. J. F., and Duncan, C. (2016). How do we want satellite remote sensing to support biodiversity conservation globally? *Methods Ecol. Evol.*, 7(6):656–665.

[Pettorelli et al., 2014b] Pettorelli, N., Safi, K., and Turner, W. (2014b). Satellite remote sensing, biodiversity research and conservation of the future. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 369(1643):20130190.

[Phillips et al., 2006] Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Modell.*, 190(3):231–259.

[Phillips and Dudík, 2008] Phillips, S. J. and Dudík, M. (2008). Modeling of species distributions with maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2):161–175.

[Phillips et al., 2009] Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., and Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol. Appl.*, 19(1):181–197.

[Pohl and Van Genderen, 1998] Pohl, C. and Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: Concepts, methods and applications. *Int. J. Remote Sens.*, 19(5):823–854.

[Ramankutty et al., 2008] Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A. (2008). Farming the planet: 1. geographic distribution of global agricultural lands in the year 2000: GLOBAL AGRICULTURAL LANDS IN 2000. *Global Biogeochemical Cycles*, 22(1).

[Ramirez-Reyes et al., 2019] Ramirez-Reyes, C., Brauman, K. A., Chaplin-Kramer, R., Galford, G. L., Adamo, S. B., Anderson, C. B., Anderson, C., Allington, G. R. H., Bagstad, K. J., Coe, M. T., Cord, A. F., Dee, L. E., Gould, R. K., Jain, M., Kowal, V. A., Muller-Karger, F. E., Norriss, J., Potapov, P., Qiu, J., Rieb, J. T., Robinson, B. E., Samberg, L. H., Singh, N., Szeto, S. H., Voigt, B., Watson, K., and Wright, T. M. (2019). Reimagining the potential of earth observations for ecosystem service assessments. *Sci. Total Environ.*, 665:1053–1063.

[Rao et al., 2020] Rao, K., Williams, A. P., Flefil, J. F., and Konings, A. G. (2020). SAR-enhanced mapping of live fuel moisture content. *Remote Sens. Environ.*, 245:111797.

[Ribot and Peluso, 2009] Ribot, J. C. and Peluso, N. L. (2009). A theory of access. *Rural Sociology*, 68(2):153–181.

[Rivera et al., 2013] Rivera, J. P., Verrelst, J., Leonenko, G., and Moreno, J. (2013). Multiple cost functions and regularization options for improved retrieval of leaf chlorophyll content and LAI through inversion of the PROSAIL model. *Remote Sensing*, 5(7):3280–3304.

[Rocchini, 2007] Rocchini, D. (2007). Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. *Remote Sens. Environ.*, 111(4):423–434.

[Rocchini et al., 2010] 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., and Neteler, M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecol. Inform.*, 5(5):318–329.

[Rochedo et al., 2018] Rochedo, P. R. R., Soares-Filho, B., Schaeffer, R., Viola, E., Szklo, A., Lucena, A. F. P., Koberle, A., Davis, J. L., Rajão, R., and Rathmann, R. (2018). The threat of political bargaining to climate mitigation in brazil. *Nat. Clim. Chang.*, 8(8):695–698.

[Rodríguez et al., 2006] Rodríguez, J. J., Kuncheva, L. I., and Alonso, C. J. (2006). Rotation forest: A new classifier ensemble method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(10):1619–1630.

[Rosenzweig, 1995] Rosenzweig, M. L. (1995). *Species Diversity in Space and Time*. Cambridge University Press.

[Rosenzweig and Abramsky, 1993] Rosenzweig, M. L. and Abramsky, Z. (1993). How are diversity and productivity related. *Species diversity in ecological communities: historical and geographical perspectives*, pages 52–65.

[Rossi et al., 1999] Rossi, G. C., Pascual, N. T., and Krsticevic, F. J. (1999). First record of aedes albopictus (skuse) from argentina. *J. Am. Mosq. Control Assoc.*, 15(3):422.

[Roughgarden et al., 1991] Roughgarden, J., Running, S. W., and Matson, P. A. (1991). What does remote sensing do for ecology? *Ecology*, 72(6):1918–1922.

[Ryan et al., 2019] Ryan, S. J., Carlson, C. J., Mordecai, E. A., and Johnson, L. R. (2019). Global expansion and redistribution of aedes-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.*, 13(3):e0007213.

[Ryan et al., 2020] Ryan, S. J., Carlson, C. J., Tesla, B., Bonds, M. H., Ngonghala, C. N., Mordecai, E. A., Johnson, L. R., and Murdock, C. C. (2020). Warming temperatures could expose more than 1.3 billion new people to zika virus risk by 2050. *Glob. Chang. Biol.*

[Saatchi et al., 2008] Saatchi, S., Buermann, W., ter Steege, H., Mori, S., and Smith, T. B. (2008). Modeling distribution of amazonian tree species and diversity using remote sensing measurements. *Remote Sens. Environ.*, 112(5):2000–2017.

[Saatchi et al., 2007] Saatchi, S. S., Houghton, R. A., Dos Santos Alvalá, R. C., Soares, J. V., and Yu, Y. (2007). Distribution of aboveground live biomass in the amazon basin. *Glob. Chang. Biol.*, 13(4):816–837.

[Santos, 2003] Santos, R. L. C. d. (2003). Updating of the distribution of aedes albopictus in brazil (1997-2002). *Rev. Saúde Pública*, 37:671–673.

[Scholes et al., 2008] Scholes, R. J., Mace, G. M., Turner, W., Geller, G. N., Jurgens, N., Larigauderie, A., Muchoney, D., Walther, B. A., and Mooney, H. A. (2008). Ecology. toward a global biodiversity observing system. *Science*, 321(5892):1044–1045.

[Scholes et al., 2012] Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H. R., Tuama, É. Ó., Faith, D. P., Mooney, H. A., Ferrier, S., Jongman, R. H. G., Harrison, I. J., Yahara, T., Pereira, H. M., Larigauderie, A., and Geller, G. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, 4(1):139–146.

[Schulte to Bühne and Pettorelli, 2018] Schulte to Bühne, H. and Pettorelli, N. (2018). Better together: Integrating and fusing multispectral and radar satellite imagery to inform biodiversity monitoring, ecological research and conservation science. *Methods Ecol. Evol.*, 9(4):849–865.

[Segurado et al., 2006] Segurado, P., Araujo, M. B., and Kunin, W. E. (2006). Consequences of spatial autocorrelation for niche-based models. *J. Appl. Ecol.*, 43(3):433–444.

[Seo et al., 2009] Seo, C., Thorne, J. H., Hannah, L., and Thuiller, W. (2009). Scale effects in species distribution models: implications for conservation planning under climate change. *Biol. Lett.*, 5(1):39–43.

[Serbin et al., 2014] Serbin, S. P., Singh, A., McNeil, B. E., Kingdon, C. C., and Townsend, P. A. (2014). Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecol. Appl.*, 24(7):1651–1669.

[Sexton et al., 2016] Sexton, J. O., Noojipady, P., Song, X.-P., Feng, M., Song, D.-X., Kim, D.-H., Anand, A., Huang, C., Channan, S., Pimm, S. L., and Townshend, J. R. (2016). Conservation policy and the measurement of forests. *Nat. Clim. Chang.*, 6(2):192–196.

[Sexton et al., 2013] Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K. M., Channan, S., DiMiceli, C., and Townshend, J. R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth*, 6(5):427–448.

[Shepard et al., 2011] Shepard, D. S., Coudeville, L., Halasa, Y. A., Zambrano, B., and Dayan, G. H. (2011). Economic impact of dengue illness in the americas. *Am. J. Trop. Med. Hyg.*, 84(2):200–207.

[Shimada et al., 2014] Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., and Lucas, R. (2014). New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.*, 155(Supplement C):13–31.

[Shipley et al., 2006] Shipley, B., Vile, D., and Garnier, E. (2006). From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science*, 314(5800):812–814.

[Shocket et al., 2018] Shocket, M. S., Ryan, S. J., and Mordecai, E. A. (2018). Temperature explains broad patterns of ross river virus transmission. *Elife*, 7.

[Siefert et al., 2015] Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W., Baraloto, C., Carlucci, M. B., Cianciaruso, M. V., and Others (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecol. Lett.*, 18(12):1406–1419.

[Sims and Gamon, 2002] Sims, D. A. and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.*, 81(2):337–354.

[Skole and Tucker, 1993] Skole, D. and Tucker, C. (1993). Tropical deforestation and habitat fragmentation in the amazon: satellite data from 1978 to 1988. *Science*, 260(5116):1905–1910.

[Soberon and Peterson, 2005] Soberon, J. and Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*.

[Soberon and Townsend Peterson, 2005] Soberon, J. and Townsend Peterson, A. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Inf.*, 2(0).

[Song et al., 2018] Song, X.-P., Hansen, M. C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E. F., and Townshend, J. R. (2018). Global land change from 1982 to 2016. *Nature*.

[Springmann et al., 2018] Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., de Vries, W., Vermeulen, S. J., Herrero, M., Carlson, K. M., Jonell, M., Troell, M., DeClerck, F., Gordon, L. J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., Godfray, H. C. J., Tilman, D., Rockström, J., and Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature*, 562(7728):519–525.

[Stanaway et al., 2016] Stanaway, J. D., Shepard, D. S., Undurraga, E. A., Halasa, Y. A., Coffeng, L. E., Brady, O. J., Hay, S. I., Bedi, N., Bensenor, I. M., Castañeda-Orjuela, C. A., Chuang, T.-W., Gibney, K. B., Memish, Z. A., Rafay, A., Ukwaja, K. N., Yonemoto, N., and Murray, C. J. L. (2016). The global burden of dengue: an analysis from the global burden of disease study 2013. *Lancet Infect. Dis.*, 16(6):712–723.

[Steven J. Phillips, Miroslav Dudík, Robert E. Schapire, 2017] Steven J. Phillips, Miroslav Dudík, Robert E. Schapire (2017). Maxent software for modeling species niches and distributions. `https://biodiversityinformatics.amnh.org/open_source/maxent/`. Accessed: 2019-5-2.

[Stuffler et al., 2007] Stuffler, T., Kaufmann, C., Hofer, S., Förster, K. P., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penné, B., Benz, U., and Haydn, R. (2007). The EnMAP hyperspectral imager—an advanced optical payload for future applications in earth observation programmes. *Acta Astronaut.*, 61(1):115–120.

[Tapia-Conyer et al., 2012] Tapia-Conyer, R., Betancourt-Cravioto, M., and Méndez-Galván, J. (2012). Dengue: an escalating public health problem in latin america. *Paediatr. Int. Child Health*, 32 Suppl 1:14–17.

[Tatem, 2017] Tatem, A. J. (2017). WorldPop, open data for spatial demography. *Sci Data*, 4:170004.

[Tatem et al., 2006] Tatem, A. J., Hay, S. I., and Rogers, D. J. (2006). Global traffic and disease vector dispersal. *Proc. Natl. Acad. Sci. U. S. A.*, 103(16):6242–6247.

[Taylor et al., 2015] Taylor, P., Asner, G., Dahlin, K., Anderson, C., Knapp, D., Martin, R., Mascaro, J., Chazdon, R., Cole, R., Wanek, W., Hofhansl, F., Malavassi, E., Vilchez-Alvarado, B., and Townsend, A. (2015). Landscape-Scale controls on aboveground forest carbon stocks on the osa peninsula, costa rica. *PLoS One*, 10(6):e0126748.

[Thuiller et al., 2009] Thuiller, W., Lafourcade, B., Engler, R., and Araújo, M. B. (2009). BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32(3):369–373.

[Thuiller et al., 2015] Thuiller, W., Pollock, L. J., Gueguen, M., and Münkemüller, T. (2015). From species distributions to meta-communities. *Ecol. Lett.*, 18(12):1321–1328.

[Tjaden et al., 2018] Tjaden, N. B., Caminade, C., Beierkuhnlein, C., and Thomas, S. M. (2018). Mosquito-Borne diseases: Advances in modelling Climate-Change impacts. *Trends Parasitol.*, 34(3):227–245.

[Tjaden et al., 2017] Tjaden, N. B., Suk, J. E., Fischer, D., Thomas, S. M., Beierkuhnlein, C., and Semenza, J. C. (2017). Modelling the effects of global climate change on chikungunya transmission in the 21st century. *Sci. Rep.*, 7(1):3813.

[Townsend et al., 2007] Townsend, A. R., Cleveland, C. C., Asner, G. P., and Bustamante, M. M. C. (2007). Controls over foliar N:P ratios in tropical rain forests. *Ecology*, 88(1):107–118.

[Troyo et al., 2009] Troyo, A., Fuller, D. O., Calderón-Arguedas, O., Solano, M. E., and Beier, J. C. (2009). Urban structure and dengue fever in puntarenas, costa rica. *Singap. J. Trop. Geogr.*, 30(2):265–282.

[Tsuda et al., 2006] Tsuda, Y., Suwonkerd, W., Chawprom, S., Prajakwong, S., and Takagi, M. (2006). Different spatial distribution of aedes aegypti and aedes albopictus along an urban-rural gradient and the relating environmental factors examined in three villages in northern thailand. *J. Am. Mosq. Control Assoc.*, 22(2):222–228.

[Tucker Lima et al., 2017] Tucker Lima, J. M., Vittor, A., Rifai, S., and Valle, D. (2017). Does deforestation promote or inhibit malaria transmission in the amazon? a systematic literature review and critical appraisal of current evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 372(1722):20160125.

[Turner, 2016] Turner, M. D. (2016). Political ecology and scale. In Richardson, D., Castree, N., Goodchild, M. F., Kobayashi, A., Liu, W., and Marston, R. A., editors, *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, volume 36, pages 1–9. John Wiley & Sons, Ltd, Oxford, UK.

[Turner, 2014] Turner, W. (2014). Conservation. sensing biodiversity. *Science*, 346(6207):301–302.

[Turner et al., 2003] Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.*, 18(6):306–314.

[UN Population Division, 2019] UN Population Division (2019). World urbanization prospects: 2018 revision. `https://population.un.org/wup/`. Accessed: 2019-5-6.

[U.S. Embassy Lima, 2020] U.S. Embassy Lima (2020). Health alert – dengue fever in peru (february 14, 2020) — U.S. embassy in peru. `https://pe.usembassy.gov/health-alert-dengue-fever-peru-february-2020/`. Accessed: 2020-3-17.

[Vanwambeke et al., 2007] Vanwambeke, S. O., Somboon, P., Harbach, R. E., Isenstadt, M., Lambin, E. F., Walton, C., and Butlin, R. K. (2007). Landscape and land cover factors influence the presence of aedes and anopheles larvae. *J. Med. Entomol.*, 44(1):133–144.

[Vellend et al., 2013] Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beauséjour, R., Brown, C. D., De Frenne, P., Verheyen, K., and Wipf, S. (2013). Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences*, 110(48):19456–19459.

[Violle et al., 2012] Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., Jung, V., and Messier, J. (2012). The return of the variance: intraspecific variability in community ecology. *Trends Ecol. Evol.*, 27(4):244–252.

[Wagman et al., 2013] Wagman, J., Grieco, J. P., King, R., Briceño, I., Bautista, K., Polanco, J., Pecor, J., and Achee, N. L. (2013). First record and demonstration of a southward expansion of aedes albopictus into orange walk town, belize, central america. *J. Am. Mosq. Control Assoc.*, 29(4):380–382.

[Waldron et al., 2017] Waldron, A., Miller, D. C., Redding, D., Mooers, A., Kuhn, T. S., Nibbelink, N., Roberts, J. T., Tobias, J. A., and Gittleman, J. L. (2017). Reductions in global biodiversity loss predicted from conservation spending. *Nature*, 551(7680):364–367.

[Walker et al., 2010] Walker, W. S., Stickler, C. M., Kellndorfer, J. M., Kirsch, K. M., and Nepstad, D. C. (2010). Large-Area classification and mapping of forest and land cover in the brazilian amazon: A comparative analysis of ALOS/PALSAR and landsat data sources. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3(4):594–604.

[Wang et al., 2010] Wang, K., Franklin, S. E., Guo, X., and Cattet, M. (2010). Remote sensing of ecology, biodiversity and conservation: a review from the perspective of remote sensing specialists. *Sensors*, 10(11):9647–9667.

[Wang et al., 2012] Wang, S. D., Miao, L. L., and Peng, G. X. (2012). An improved algorithm for forest fire detection using HJ data. *Procedia Environmental Sciences*, 13(Supplement C):140–150.

[Waring and Running, 2010] Waring, R. H. and Running, S. W. (2010). *Forest Ecosystems: Analysis at Multiple Scales*. Elsevier.

[Whittaker et al., 2005] Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M., and Willis, K. J. (2005). Conservation biogeography: assessment and prospect: Conservation biogeography. *Diversity and Distributions*, 11(1):3–23.

[Whittaker et al., 2001] Whittaker, R. J., Willis, K. J., and Field, R. (2001). Scale and species richness: towards a general, hierarchical theory of species diversity: Towards a general theory of diversity. *J. Biogeogr.*, 28(4):453–470.

[Wiens, 1989] Wiens, J. A. (1989). Spatial scaling in ecology. *Funct. Ecol.*, 3(4):385–397.

[Wiens et al., 2009] Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., and Snyder, M. A. (2009). Niches, models, and climate change: assessing the assumptions and uncertainties. *Proc. Natl. Acad. Sci. U. S. A.*, 106 Suppl 2:19729–19736.

[Wiens et al., 2010] Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., Damschen, E. I., Jonathan Davies, T., Grytnes, J.-A., Harrison, S. P., Hawkins, B. A., Holt, R. D., McCain, C. M., and Stephens, P. R. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.*, 13(10):1310–1324.

[Wilson and Jetz, 2016] Wilson, A. M. and Jetz, W. (2016). Remotely sensed High-Resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLoS Biol.*, 14(3):e1002415.

[Withers and Meentemeyer, 1999] Withers, M. A. and Meentemeyer, V. (1999). Concepts of scale in landscape ecology. In Klopatek, J. M. and Gardner, R. H., editors, *Landscape Ecological Analysis: Issues and Applications*, pages 205–252. Springer New York, New York, NY.

[Wolford et al., 2013] Wolford, W., Borras, Jr., S. M., Hall, R., Scoones, I., and White, B. (2013). Governing global land deals: The role of the state in the rush for land. *Dev. Change*, 44(2):189–210.

[Wright et al., 2004] Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets,

U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G., Veneklaas, E. J., and Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428(6985):821–827.

[Yao et al., 2015] Yao, W., van Leeuwen, M., Romanczyk, P., Kelbe, D., and van Aardt, J. (2015). Assessing the impact of sub-pixel vegetation structure on imaging spectroscopy via simulation. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*, volume 9472, page 94721K. International Society for Optics and Photonics.

[Yates et al., 2018] Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., Heikkinen, R. K., Heinänen, S., Jones, A. R., Krishnakumar, P. K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M. B., Moreno-Amat, E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, G., Peterson, A. T., Rapacciuolo, G., Roberts, J. J., Ross, R. E., Scales, K. L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L. G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M. J., Zharikov, Y., Zurell, D., and Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends Ecol. Evol.*, 33(10):790–802.

[Ye and Glantz, 2005] Ye, Q. and Glantz, M. H. (2005). The 1998 yangtze floods: The use of Short-Term forecasts in the context of seasonal to interannual water resource management. *Mitigation and Adaptation Strategies for Global Change*, 10(1):159–182.

[Zhu et al., 2012] Zhu, Z., Woodcock, C. E., and Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available landsat imagery. *Remote sensing of environment*, 122:75–91.