Modeling tropical montane forest biodiversity

The potential of multispectral remote sensing

Doctoral thesis
by Christine I. B. Wallis
Modeling tropical montane forest biodiversity –
The potential of multispectral remote sensing

Dissertation
zur
Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Geographie
der Philipps-Universität Marburg
vorgelegt von

Christine Isabeau Bernarde Wallis
aus Tecklenburg

Marburg an der Lahn, September 2018

Diese Arbeit steht unter der
Creative Commons by-nc-sa-Lizenz
Cover image: Tropical mountain rainforest and German research station (Estacion Científica San Francisco) of the San Francisco valley at approx. 2000 m a.s.l. (Sigtierras orthophoto 2011)

Erstgutachter: Prof. Dr. Jörg Bendix
Zweitgutachter: Prof. Dr. Roland Brandl

"In the end we will conserve only what we love, we will love only what we understand, and we will understand only what we are taught."

Baba Dioum
# Contents

1 **General introduction** ........................................... 1  
1.1 Motivation ......................................................... 3  
1.1.1 Remote sensing applications to biodiversity ................. 3  
1.1.2 Drivers of biodiversity and their remotely sensed proxies .... 4  
1.1.3 Potential and challenges in modeling tropical montane forest biodiversity using multispectral data ......... 6  
1.2 Conceptual design ............................................... 10  
1.2.1 Aims, objectives and hypotheses ............................ 10  
1.2.2 Structure of this thesis ................................... 10  
1.2.3 Study area and data acquisition ............................ 12  
References ............................................................ 15  

2 **Contrasting performance of Lidar and optical texture models in predicting avian diversity in a tropical mountain forest** 25  
Abstract ...................................................................... 27  
2.1 Introduction ......................................................... 27  
2.2 Data and Methods ............................................... 29  
2.2.1 Study area ....................................................... 29  
2.2.2 Avian field data and proxies for avian diversity .......... 31  
2.2.3 Remote sensing indicators ................................. 32  
2.2.4 Analyses ......................................................... 33  
2.3 Results .................................................................. 35  
2.3.1 Diversity proxies .............................................. 35  
2.3.2 PLS regression model comparison ......................... 35  
2.3.3 Feature selection .............................................. 36  
2.3.4 Prediction maps bird community ........................... 38  
2.4 Discussion .......................................................... 38  
2.4.1 Forest structure and spectral respond ....................... 38  
2.4.2 Shannon diversity ............................................ 39  
2.4.3 Phylodiversity ............................................... 39  
2.4.4 Bird community .............................................. 40  
2.4.5 Limitations and recommendations ......................... 41  
2.4.6 Conclusions and conservation implications ............... 42  
Acknowledgements ....................................................... 42  
References ............................................................ 42  

3 **Remote sensing improves prediction of tropical montane species diversity but performance differs among taxa** 53  
Abstract ...................................................................... 55  
3.1 Introduction ......................................................... 55  
3.2 Data and Methods ............................................... 57  
3.2.1 Study area ....................................................... 57
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B</strong> Chapter 3</td>
<td>143</td>
</tr>
<tr>
<td><strong>C</strong> Chapter 4</td>
<td>151</td>
</tr>
<tr>
<td><strong>D</strong> Additional information on the chapters 2-4 (not published)</td>
<td>161</td>
</tr>
<tr>
<td><strong>Erklärung</strong></td>
<td>169</td>
</tr>
</tbody>
</table>
Chapter 1

General introduction
1.1 Motivation

Biodiversity describes “the diversity of genes, populations, species, communities, and ecosystems” (Hassan et al., 2005). As proposed by Franklin (1988) and Noss (1990), structural, compositional (or taxonomical), and functional diversity are the three main aspects of biodiversity. These aspects comprise different levels of biological organization: regional landscape, community–ecosystem, population–species, and genetics (Noss, 1990). Structural diversity is related to structural patterns such as habitat complexity within a community or number of patches at the landscape scale. Taxonomic diversity describes the variety of species by using, for example, the simple number of species, indices of species or genetic diversity. Functional diversity faces ecological processes (e.g., competition, herbivory, and photosynthesis), ecosystem processes (e.g., primary production, decomposition, and evapotranspiration) and ecosystem functions (e.g., nutrient regulation, and water supply; Lovett et al., 2005; Martinez, 1996; Pettorelli et al., 2017). The classification of Noss (1990) has been extended over time; e.g., with an additional category for phylodiversity which relates to taxonomic diversity and functional diversity (see e.g., Lausch et al., 2018).

Biodiversity, thus, encompasses not only the often investigated species diversity but also ecological and ecosystem processes and functions at different levels of complexity. There is also growing evidence that the functioning of ecosystems provides services of high value for human well-being (Cardinale et al., 2012), e.g., raw material production, carbon sequestration, recreational experience, and food production (Costanza et al., 1997; Petter et al., 2013).

The ongoing human-induced land-use modifications, however, affect biodiversity and are a risk to the provision of ecosystem services (Hanski, 2011). Biodiversity loss has been identified as a core planetary boundary that has already exceeded its threshold triggering nonlinear, abrupt environmental change (Rockström et al., 2009; Steffen et al., 2015). To diminish biodiversity loss, the Convention on Biological Diversity (CBD) has set up 20 targets divided into five strategic goals to be achieved by 2020 (CBD, 2014). These targets provide a framework to take effective and urgent action to preserve biodiversity at global, regional, and national scales (Han et al., 2014). To track the progress of these goals, robust monitoring systems are necessary (CBD, 2014). Two of the Aichi targets emphasize the importance of protecting and safeguarding ecosystems that are of particular relevance for biodiversity and ecosystem services (target 11 and 14; CBD, 2014).

Tropical mountain forests, particularly Andean rainforests, are among the most diverse, but at the same time most threatened biodiversity hotspots in the world (Tapia-Armijos et al., 2015). Their outstanding biodiversity mediates diverse ecosystem functions and provides essential ecosystem services to humans from the regional scale (e.g., water supply) to the global scale (e.g., carbon sequestration; Curatola Fernández et al., 2015; Hooper et al., 2005). Therefore, monitoring biodiversity in these ecosystems is necessary. But due to the inaccessibility of most montane regions and exacerbated by the high species abundance and richness of tropical rainforests, an acceptable field-sampling of taxonomic, structural and functional diversity would demand far more resources...
1.1 Motivation

than available. Spatially explicit information about such diverse areas are therefore needed to provide effective indicator systems for an area-wide monitoring of biodiversity.

Remote sensing and geoinformation sciences offer great opportunities to derive comprehensive environmental indicators for characterizing ecosystems in areas where an effective field-sampling is hardly possible due to their inaccessibility and size (Pettorelli et al., 2015). Such indicators can be used to model biodiversity spatially and temporally (Pettorelli et al., 2014). Basically there are two types of remote sensing sensors: active and passive sensors. Whereas active sensors send their own signal to Earth and receive the backscattered reflection, passive sensors depend on the reflection of sunlight from the Earth’s surface (Prasad et al., 2015). Both sensor types are available in different spatial and temporal resolutions and differ in terms of their availability from ground, airborne and spaceborne platforms. The current strengths and limitations of remote sensing therefore primarily concern the sensor properties.

Although remote sensing has improved very quickly and new sensors offer new possibilities, there are still weaknesses in the reproduction of environmental properties. These are primarily related to sensor limitations in terms of their availability, their costs or their technical challenges (Anderson et al., 2016; Nagendra et al., 2013; Popescu et al., 2011; Wolter et al., 2009). To date, only satelliteborne multispectral sensors at low cost with a moderate to high repetition rate and ground coverage provide information to derive spatial indicators suitable for an operational biodiversity monitoring system (Pettorelli et al., 2016). These sensors are often used to model biological entities and attributes at continental to global scales. Since these models often neglect within-habitat heterogeneity, the high biodiversity in hotspot areas such as the tropical montane rainforests is often underestimated (Foody, 2003). To enhance multispectral models of tropical montane biodiversity, the development of comprehensive indicators at the landscape scale is therefore necessary.

1.1.1 Remote sensing applications to biodiversity

Most remote sensing studies are based on empirical relationships between remotely sensed environmental predictors and biodiversity variables derived from biological in situ observations. These in situ observations might be related to biological entities or attributes of interest in biodiversity monitoring and are used to train and validate biodiversity models (also known as ground truthing; Ferrier et al., 2017).

Since biodiversity is a multidimensional concept, remote sensing-driven studies need to face the multiple aspects of biodiversity and levels of biological organization in space and time. However, in megadiverse ecosystems, the high taxonomic, functional and structural diversity cannot be assessed continuously, and spatially explicit at the species-level. Therefore, at the landscape scale, remote sensing-driven studies often use community-level approaches without providing explicit information on the individual species comprising this diversity (Ferrier et al., 2017). To
prioritize among different biodiversity measures, essential biodiversity variables (EBV) have been proposed (Pereira et al., 2013). Skidmore et al. (2015) suggested ten important EBV that can be tracked from space considering species populations, species traits, ecosystem structure and ecosystem functions. These EBV are, to date, unequally investigated in remote sensing studies. Certain measures of taxonomic and functional diversity such as species richness, habitat structure, and primary production, have often been investigated using remote sensing data (Warren et al., 2014; Zhao and Running, 2008). Less long studied by remote sensing but recently investigated biodiversity variables are phylogenetic diversity, species turnover as well as functional leaf traits (He et al., 2015; McManus et al., 2016; Pettorelli et al., 2017). The success of remote sensing data in estimating different biodiversity variables can vary considerably, depending on the biological entities and attributes studied. This is particularly important for the development of remotely sensed indicator systems of biodiversity. A deeper understanding of these different predictabilities of biological diversity is, thus, urgently needed, especially in biota, where biodiversity inventories are difficult to obtain.

The opportunities to model biodiversity using remotely sensed environmental proxies are manifold. Recent advantages in sensor improvements have foster the use of airborne and spaceborne platforms equipped with Light detection and ranging (Lidar) and hyperspectral sensors. Lidar data, based on active sensors, has often been used to model the vertical structure of plants for habitat characterization (Vierling et al., 2008; Wulder et al., 2008; Zimble et al., 2003). Lidar sensors are independent of cloud cover and can therefore be used in regions such as tropical cloud forests. In addition, hyperspectral sensors from airborne platforms are promising for modeling biodiversity variables such as plant species diversity (Peng et al., 2018). The use of these sensors is, however, limited due to their costs, their availability and their technical challenges (Anderson et al., 2016; Nagendra et al., 2013; Popescu et al., 2011; Wolter et al., 2009). For this reason, to date, the real strengths of remote sensing for monitoring biodiversity rely on spaceborne multispectral sensors. Satellite data offer a reasonable lifetime and repetition rate for monitoring biodiversity. There are a number of commercial satellite sensors with suitable combinations of spectral, spatial and temporal resolution (e.g., IKONOS, Quickbird). A large part of remote sensing studies is, however, based on freely available data, for example, provided by the Earth Observing System (EOS) missions (e.g., Landsat and Moderate Resolution Imaging Spectroradiometer, MODIS, images; Popkin, 2018). Developing countries with limited access to expensive remote sensing data particularly benefit from such open data archives. But the corresponding sensors come along with limitations in terms of either their spectral or spatial resolutions. In tropical montane ecosystems, furthermore, immense cloud cover makes it difficult to collect cloudless satellite scenes. To circumvent the problem of cloud cover, airborne orthophotos with a high spatial but low spectral resolution (either visible or multispectral wavelength bands) are often used. Consequently, yet, the ability to
model and map certain aspects of biodiversity at the landscape scale using multispectral data is limited (Homolová et al., 2013; Martínez et al., 2016; Ollinger, 2011).

1.1.2 Drivers of biodiversity and their remotely sensed proxies

The choice of remotely sensed predictors in biodiversity modeling is usually based on known theory, relationships, processes and rules. MacArthur (1984), for example, proposed that biodiversity is a function of climatic stability, vegetation productivity, and habitat structure. Many studies have investigated the link between these environmental properties and different aspects of biodiversity (e.g., Gaston, 2000; Rosenzweig, 1995). Additional hypotheses have been used to explain patterns of biodiversity gradients in particular (e.g., the species-energy hypothesis, the metabolic niche hypothesis, physiological tolerance hypothesis; Currie et al., 2004; Clarke, 2006; Currie, 1991). However, most of these hypotheses again relate to environmental filters such as temperature, vegetation productivity or habitat structure. In recent decades, remotely sensed proxies of these environmental drivers have been used to examine their effects on taxonomic, functional and structural diversity (Turner et al., 2003; Wu and Liang, 2018).

Particularly along elevational gradients, elevation has been used as a simple proxy for changes in species diversity based on the elevational lapse of air temperature change (Kübler et al., 2016; Malsch et al., 2008; Nakamura et al., 2015). As elevation can be derived from free global digital elevation model (DEM) data sources, its use as a remotely sensed proxy for the spatial variation of temperature is widely accepted. Peters et al. (2016) revealed that temperature is the main driver of elevational species diversity at the multi-taxa community level along an elevation gradient on Mt. Kilimanjaro. For single taxa, however, they found varying environmental drivers. The results are strongly taxon-dependent, since the temperature dependence varies greatly among taxonomic groups (e.g., Fiedler et al., 2008; Tiede et al., 2016). However, the abiotic conditions in forests also change with elevation (McCain and Grytnes, 2010). Elevation and its derivatives, slope and topographic position, therefore, surrogate spatial variation in nutrient supply, wind exposure, and soil hydrology in addition to temperature (Wilcke et al., 2008; Bendix et al., 2008; Wagemann et al., 2015).

Vegetation productivity provides resources such as the availability of food, which lead to a higher abundance of individual species (Evans et al., 2005; Wright, 1983). In turn, according to the ‘more individuals hypothesis’ a higher number of individuals is related to a higher number of species that can have viable populations (Gaston, 2000). Proxies of habitat productivity are most often derived from spectral indices such as the Normalized-Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI) of multispectral remote sensing sensors (Glenn et al., 2008). These vegetation indices are based on the spectral reflection of leaf pigments. Generally healthy vegetation absorbs light in the visible spectrum and reflects in the near infrared spectrum (Bannari et al., 1995). In tropical mountain rainforests, however, elevation might affect vegetation
and its spectral respond at the sensor. Therefore common vegetation indices might not explain complementary information to elevation in modeling biodiversity.

A higher habitat structure, generally, provides more potential niches and different resources allowing a higher number of species to coexist (Rosenzweig, 1995). It should be noted that there are no clear boundaries between structural diversity (as biodiversity variable) and habitat structure (as remotely sensed predictor). Although, in theory biodiversity is also defined by structural diversity as proposed by Franklin (1988) and Noss (1990), structural diversity of plants, particularly forests, can be used to model taxonomic and functional diversity (see also Figure 1.1a, b; Zellweger et al., 2013). Remote sensing provides suitable information to derive cost-effective spatial proxies for habitat structure and heterogeneity (Beguet et al., 2012; Chambers et al., 2007; Rocchini et al., 2010; Wood et al., 2012; Zimble et al., 2003). Topographical metrics can be used as simple proxies of habitat structure derived from digital elevation models (Turner et al., 2003). To account for the three-dimensional structure of vegetation, active remote sensing sensors such as synthetic aperture radar (SAR) and Lidar have been used also (González-Jaramillo et al., 2018). Lidar metrics provide more complex remote sensing products that surrogate vertical vegetation structure (Vierling et al., 2008). Lidar metrics have been been used, for example, to model the relationship between canopy structure and (structural) plant traits such as the leaf area index (LAI), tree density and canopy height (Martínez et al., 2016; Zhao and Popescu, 2009; Ferraz et al., 2016). However, in tropical regions, Lidar surveys are known to be associated with serious problems due to the dense canopy structure, which leads to low model performance (Müller et al., 2010). A recent study also suggested to use hyperspectral data for modeling plant species diversity due to its high spectral variation (Peng et al., 2018). A technique to derive the horizontal habitat structure derived from multispectral remote sensing is provided by the image textural approach. Image textures derived from optical images characterize the spatial variation of spectral values within a defined environment (fixed or moving window; Haralick, 1979; Wood et al., 2012). Within this defined window, the central pixel is recalculated based on statistics applied on all pixels in that window (see also Figure 1.1b). Texture, therefore, is a measure for the spatial arrangement of the spectral values and its metrics have been successfully used as proxies of habitat structure (Wood et al., 2012a) and to characterize biomass (Lu, 2005) and species diversity (St-Louis et al., 2009; Wood et al., 2013).

1.1.3 Potential and challenges in modeling tropical montane forest biodiversity using multispectral data

Monitoring tropical montane biodiversity needs effective spatially explicit indicators of biodiversity. Owing to the high costs of airborne Lidar and hyperspectral surveys, access to their data is severely restricted and often unaffordable for nonprofit organizations. Consequently, the environmental proxies derived from multispectral remote sensing data need to be improved.
Figure 1.1: Pathways to spatially model and map tropical montane biodiversity variables at the community-level using multispectral remote sensing data. a) Sampling of in situ data of interest for biodiversity monitoring; these could be related to different aspects of biodiversity, e.g., taxonomic, functional and structural diversity (Noss, 1990; Franklin, 1988). b) To surrrogate environmental drivers of biodiversity (such as temperature, vegetation productivity, and habitat structure) three remote sensing predictor sets are proposed: topographical, spectral and textural metrics. Textural metrics take the spatial variation of spectral metrics into account. Within the image textural approach the centered pixel (red) is recalculated based on texture statistics within a defined neighborhood (here a 3 pixel \times 3 pixel window). Black arrows define the movement of the moving window (here in two directions); textures will be calculated for each shift. c) Statistical methods are necessary to calibrate and validate remote sensing predictors using, for example regression models, considering the constraints of in situ data (e.g., a low number of observations). d) The spatial predictors can subsequently be used to predict the derived model(s) across space.
The pathways to model and map biodiversity variables with multispectral data are illustrated in Figure 1.1. A remote sensing indicator system for tropical montane rainforests needs to address multiple biodiversity variables considering taxonomic, functional and structural diversity at the various spatial and biological scales, but in particular at the community-level (Figure 1.1a). Related to the classification of MacArthur (1984), remotely sensed environmental proxies need to address environmental drivers of biodiversity such as temperature, vegetation productivity, and habitat structure (Figure 1.1b). To enhance predictions of biodiversity using multispectral data, remote sensing predictors need to be improved. Such an improvement might be achieved by the further development of spectral metrics such as vegetation indices and textural metrics as proxies for the habitat structure. However, their benefit among different variables of tropical montane biodiversity is not well explored yet. In addition, remote sensing-driven models need to address the constraints of in situ observations. As the sampling of in situ data is labor-intensive, the number of samples is often low. Linking remote sensing data with in situ observations, thus, involves a low number of sampling points and has to address this challenge using appropriate statistical models (Figure 1.1c). Partial least squares regression, for example, is able to treat a higher number of predictor variables than observations (number of samples) and strong colinearity among predictors (Carrascal et al., 2009). The development of remotely sensed environmental proxies and statistical methods would subsequently improve maps of biodiversity (Figure 1.1d).

Thus, modeling biodiversity in diverse and topographically complex rainforests is advantageous but simultaneously challenging. From the limitations, in summary, the following challenges result, which are relevant for this thesis:

- derivation of enhanced proxies of vegetation productivity and habitat structure;
- to overcome a low number of sampling points by the application of appropriate statistical methods;
- applicability of remotely sensed environmental proxies (e.g. surrogating temperature, vegetation productivity and habitat structure) among different aspects of biodiversity.
1.2 Conceptual design

1.2.1 Aims, objectives and hypotheses

It is increasingly urgent to improve operational remote sensing proxies for modeling tropical montane biodiversity. This is of particular concern for the development of a spatially explicit monitoring system of biodiversity. In this thesis, I investigated multispectral remote sensing data and biodiversity data in a diverse but also topographically complex tropical rainforest in southern Ecuador.

In line with the challenges proposed in the previous chapter, the main aims of my thesis were therefore to provide:

- improved models of biodiversity based on multispectral remote sensing data;
- a better understanding of the spatial variation of tropical montane biodiversity;
- and a better understanding of the remotely sensed environmental drivers of different biodiversity variables.

The following hypotheses have been put forward:

H1 Models of biodiversity benefit from the inclusion of
(a) textural metrics as proxies for changes in abiotic conditions along habitat structure
(b) spectral metrics as proxies for changes in abiotic conditions along vegetation productivity.

H2 The predictive performances of spectral and textural metrics differ across biodiversity variables considering taxonomic and functional diversity, various diversity indices and taxonomic groups.

My objectives were to (i) compare the performance of spectral and textural metrics in modeling taxonomic and functional diversity, different diversity indices and taxa, (ii) assess the benefit of spectral and textural metrics as complementary predictors to topographical metrics, and (iii) compare airborne and spaceborne multispectral sensors in modeling biodiversity. As a benchmark for comparison, I used topographical metrics as surrogates for changes in abiotic conditions along the elevational gradient and as simple metrics of habitat structure.

1.2.2 Structure of this thesis

To access my aims and objectives, I followed the workflow illustrated in Figure 1.2. In Chapter 2-4, I modeled biodiversity variables using topographical, spectral and textural metrics. The first two studies focused on taxonomic diversity considering different indices of species diversity. The
third study concentrated on functional diversity considering ecosystem processes and functions. According to the nature of topographical and textural metrics, all three studies considered structural diversity of plant species, even though this was not directly modeled. The three studies of Chapter 2–4 have been either published in or submitted to peer-reviewed international scientific journals. They are structured as follows:

1. In Chapter 2, I focused on the potential of image textural metrics derived from very high resolution satellite data to model biodiversity (H1a). I exemplarily used bird species records to calculate indices of avian diversity since bird species are known to be highly sensitive to habitat structure (Cody, 1981). I compared the performance of textural metrics to model
1.2 Conceptual design

the diversity of avian species with commonly used structural Lidar metrics. Since model performance may differ between the different indices of species diversity, I compared the Shannon diversity, the phylo-\(\alpha\)-diversity and a community composition representing means of species turnover (H2). Although taxonomic diversity was targeted in this chapter, one of the species diversity indices, phylo-\(\alpha\)-diversity, is also known as a proxy for functional diversity (Flynn et al., 2011; Lyashevska and Farnsworth, 2012).

2. In Chapter 3, I focused on the performance of spectral and textural metrics to model biodiversity of multiple taxonomic groups (H2). The approach of Chapter 2 was therefore generalized investigating models of species richness and species turnover of six taxonomic groups. These were trees, pyraloid moths, geometrid moths, arctiinae moths, ants and birds. I modeled their species richness and turnover (decomposed into two dimensions) as functions of topographical and textural metrics derived from airborne very high spatial resolution orthophotos. As a benchmark, I also calculated models that only used topographical metrics and models that used a combination of topographical and spectral metrics. In addition to the NDVI, a normalized vegetation index based on the red and the blue wavelength bands was used as spectral metric (H1b; Appendix B Table B6).

3. In Chapter 4, I investigated the potential of spectral and textural metrics in modeling functional diversity. I used forest biomass, productivity and morphological and biochemical canopy traits (measured as community-weighted means of functional leaf traits) to account for functional diversity. These ecosystem processes and functions were modeled using spectral and textural metrics derived from Landsat 8, an operational satellite equipped with a medium resolution multispectral sensor (H1; Vermote et al., 2016). My objective was to assess the benefit of multispectral data in comparison to topographical metrics. To do so, I tested whether topographical, spectral and textural metrics are able to explain functional diversity along the elevation gradient and within single elevation sites.

1.2.3 Study area and data acquisition

The three studies were conducted in Southern Ecuador, with a focus on the Andean cordillera between the cities of Loja und Zamora, an area of approximately 600 km\(^2\). For logistic reasons, the field-work, that forms the basis of this thesis, was concentrated on the three study sites of Bombuscaro, San Francisco and Cajanuma (Figure 1.3a, black rectangles). These three study sites followed an elevational gradient from approx. 1000 m a.s.l. in Bombuscaro to 2000 m a.s.l. in San Francisco to 3000 m a.s.l. in Cajanuma. All three study sites partly overlap with the Podocarpus Nationalpark. The Reserva Biológica de San Francisco in the vicinity of the research station (Estacion Cientifica San Francisco) is located in the northern part of the study area (Estación
Figure 1.3: a) Location of the study area in Ecuador (left) and a detailed view of the study area with a particular focus on the three black rectangles covering the study sites of Cajanuma, San Francisco, and Bombuscaro (right). The elevation data is derived from the Aster elevation model (see Table 1.2). b) Field plots of all sampled biological entities and attributes (see Table 1.1) and forest loss between 2001 and 2016 (Figure 1.3b; I reclassified forest loss using the classification of Hansen et al., 2013). All three study sites have undergone forest loss events, but the sampling plots were not directly affected and the forest losses in the vicinity have been low. Non-Forest areas were masked out based on the land cover classification of Göttlicher et al. (2009).

Científica de San Francisco, ECSF, DFG field station: 3°58’18” S and 79°4’45” W, altitude: 1860 m a.s.l.; Bendix et al., 2006).

In situ data (or ground truth data), sampled by me and former research groups, was not constantly available for all three study sites (Table 1.1). For this reason, different remote sensing sources were necessary to address the scales and requirements of the sampled biological entities and attributes (Table 1.2). While Lidar and multispectral Quickbird images were available for San Francisco only, orthorectified aerial images derived from SIGTIERRAS with a resolution of 0.3 m and four spectral bands (red, green, blue and near infra-red) were available for all three study sites. The latest data...
Table 1.1: Sampled biological in situ data within the study area. Further information can be found in Chapter 2-4.

<table>
<thead>
<tr>
<th>Biological data</th>
<th>Number of samples</th>
<th>Recording date</th>
<th>Study site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird species</td>
<td>30</td>
<td>2000-2002</td>
<td>San Francisco</td>
</tr>
<tr>
<td>Tree species</td>
<td>50</td>
<td>2007-2008</td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
<tr>
<td>Moth species</td>
<td>20</td>
<td>1999-2000</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
<tr>
<td>Ant species</td>
<td>27</td>
<td>2014</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
<tr>
<td>C/N ratio in plant tissues</td>
<td>27</td>
<td>2014</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
<tr>
<td>Functional leaf traits</td>
<td>54</td>
<td>2014</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
<tr>
<td>Annual wood production, annual fine</td>
<td>54</td>
<td>2008, 2009</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td>litter production</td>
<td></td>
<td></td>
<td>San Francisco</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cajanuma</td>
</tr>
</tbody>
</table>

originated from a cloudless Landsat 8 scene from 2016, which also covered all three study sites. Although there was a time delay between the recording date of in situ data and the recording date of the remote sensing data, I found that none of the sampling plots was directly affected by land-cover changes between 2001 and 2016 (Figure 1.3b).

I conclude that the forests in this study area as well as the data available are well suited to study the predictive performance of remotely sensed environmental predictors and to assess the potential of multispectral and textural metrics in modeling various biodiversity variables along a topographically complex and diverse ecosystem.
### Table 1.2: Remote sensing data used in this thesis.

<table>
<thead>
<tr>
<th>Remote sensing data</th>
<th>Sensor type</th>
<th>Platform</th>
<th>Spatial resolution</th>
<th>Recording date</th>
<th>Study site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Airborne laser scanning (Leica ALS-50-II laser scanner)</strong></td>
<td>Active</td>
<td>Airborne</td>
<td>1 m</td>
<td>01/03/12</td>
<td>San Francisco</td>
</tr>
<tr>
<td><strong>Quickbird</strong></td>
<td>Passive, multispectral</td>
<td>Spaceborne</td>
<td>2.6</td>
<td>10/22/10</td>
<td>San Francisco</td>
</tr>
<tr>
<td><strong>Sigtierras</strong></td>
<td>Passive, multispectral</td>
<td>Airborne</td>
<td>0.3 m</td>
<td>11/25/10</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td><strong>Sigtierras DEM</strong></td>
<td>Active</td>
<td>Airborne</td>
<td>6 m</td>
<td>11/25/10</td>
<td>San Francisco</td>
</tr>
<tr>
<td><strong>Landsat-8 OLI</strong></td>
<td>Passive, multispectral</td>
<td>Spaceborne</td>
<td>30 m</td>
<td>11/20/16</td>
<td>Bombuscaro</td>
</tr>
<tr>
<td><strong>Aster Elevation model</strong></td>
<td>Passive, optical, stereo-pair images</td>
<td>Spaceborne</td>
<td>30 m</td>
<td>10/17/11</td>
<td>San Francisco</td>
</tr>
</tbody>
</table>

### References


Carrascal, L.M., Galván, I., Gordo, O., 2009. Partial least squares regression as an alternative to cur-
References


Fiedler, K., Brehm, G., Hilt, N., Süßenbach, D., Häuser, C.L., 2008. Variation of diversity patterns across moth families along a tropical elevational gradient., in: Caldwell, M.M., Heldmaier, G.,
Chapter 1 General introduction


and Trends Working Group of the Millennium Ecosystem Assessment. Number v. 1 in The
He, K.S., Bradley, B.A., Cord, A.F., Rocchini, D., Tuanmu, M.N., Schmidtlein, S., Turner, W.,
Wegmann, M., Pettorelli, N., 2015. Will remote sensing shape the next generation of species
of optical-based remote sensing for plant trait mapping. Ecological Complexity 15, 1–16.
Kübler, D., Hildebrandt, P., Günter, S., Stimm, B., Weber, M., Mosandl, R., Muñoz, J., Cabrera, O.,
Zeilinger, J., Silva, B., 2016. Assessing the importance of topographic variables for the spatial
distribution of tree species in a tropical mountain forest. Erdkunde 70, 19–47. doi:10.3112/erdkunde.2016.01.03.
S., Mollenhauer, H., Paasche, H., Paulheim, H., Pause, M., Schweitzer, C., Schmullius, C., Settele,
Forest Health with Remote Sensing, Part III: Requirements for a Scalable Multi-Source Forest
doi:10.3390/rs10071120.
Lyashevska, O., Farnsworth, K.D., 2012. How many dimensions of biodiversity do we need?
University Press, Princeton, N.J.
An analysis of declining ant species richness with increasing elevation at Mount Kinabalu,
Sabah, Borneo. Asian Myrmecology 2, 33–49.
Martínez, O.J., Oscar, J., Premier, A.K., Günter, S., Ramos Bendaña, Z., Vierling, L., Galbraith,
Chapter 1 General introduction


Chapter 1 General introduction


Chapter 2

Contrasting performance of Lidar and optical texture models in predicting avian diversity in a tropical mountain forest

Christine I. B. Wallis, Detlev Paulsch, Brenner Silva, Giulia F. Curatola Fernández, Roland Brandl, Nina Farwig, Jörg Bendix

Abstract

Ecosystems worldwide are threatened by the increasing impact of land use and climate change. To protect their diversity and functionality, spatially explicit monitoring systems are needed. In remote areas, monitoring is difficult and recurrent field surveys are costly. By using Lidar or the more cost-effective and repetitive optical satellite data, remote sensing could provide proxies for habitat structure supporting measures for the conservation of biodiversity. Here we compared the explanatory power of both, airborne Lidar and optical satellite data in modeling the spatial distribution of biodiversity of birds across a complex tropical mountain forest ecosystem in southeastern Ecuador. We used data from field surveys of birds and chose three measures as proxies for different aspects of diversity: (i) Shannon diversity as a measure of $\alpha$-diversity that also includes the relative abundance of species, (ii) phylodiversity as a first proxy for functional diversity, and (iii) community composition as a proxy for combined $\alpha$- and $\beta$-diversity. We modeled these diversity estimates using partial least-square regression of Lidar and optical texture metrics separately and compared the models using a leave-one-out validated $R^2$ and root mean square error. Bird community information was best predicted by both remote sensing datasets, followed by Shannon diversity and phylodiversity. Our findings reveal a high potential of optical texture metrics for predicting Shannon diversity and a measure of community composition, but not for modeling phylodiversity. Generalizing from the investigated tropical mountain ecosystem, we conclude that texture information retrieved from multispectral data of operational satellite systems could replace costly airborne laser-scanning for modeling certain aspects of biodiversity.

2.1 Introduction

Modifications and losses of habitats by human impact affect biodiversity and ecosystem functionality worldwide (Cardinale et al., 2012; Newbold et al., 2015; Popp et al., 2014; Thomas et al., 2004). Thus, biodiversity monitoring systems are prerequisite to develop effective management strategies (Cardinale et al., 2012). In this context, appropriate information derived from remote sensing provides cost-effective tools that support current monitoring systems. Remote sensing has been successfully used to model species diversity of various species groups across a range of scales and habitats (Rochchini et al., 2015; Turner et al., 2003). These studies are based on the assumption that a higher environmental heterogeneity can host a higher number of species (Rosenzweig, 1995; Tews et al., 2004). Presently, only few cost-effective monitoring systems have been developed and tested for tropical biodiversity hotspots, which predict aspects of biodiversity across space and time with an extent and grain that is sufficient for meaningful management decisions.

In our work, we will concentrate on birds as indicators of biodiversity because diversity and composition of bird assemblages is highly correlated with habitat structure (Harper and Hawksworth, 1994; Lemaitre et al., 2012; Muller et al., 2010; Pearson, 1994) as eg. tree trunk density (Bergner et al., 2015). Moreover, birds respond quickly to habitat modifications, and occurrence as well as abundance of bird species can be easily sampled (Wang et al., 2014). Furthermore, their diversity is often correlated to the diversity of other taxa, such as butterflies (Blair, 1999) and plants (Kati et al., 2004). Larsen et al. (2012) found that proxies of bird diversity
2.1 Introduction

may represent overall species diversity, especially in areas where birds are more speciose than other taxa. Thus, time and costs of monitoring biodiversity might be reduced by first linking field data on birds with remotely sensed habitat structure and subsequently using the derived functions to predict biodiversity across larger areas. Until the early 21st century, bird–habitat relationships were mostly assessed using landscape metrics derived from classified satellite images across multiple habitats (Gottschalk et al., 2005). However, such studies at larger scales often ignored within-habitat heterogeneity, which may lead to more potential niches and consequently more bird species (Bar-Massada et al., 2012; Cody, 1981). For example, in the Andes of South America, present models at continental scales inadequately explain avian species-richness patterns, and new predictors are urgently needed (Rahbek et al., 2007).

Bird-habitat studies have successfully assessed vertical heterogeneity metrics derived from active sensors, such as light detection and ranging (Lidar) in the USA (Goetz et al., 2007; Lesak et al., 2011; Vierling et al., 2008; Vogeler et al., 2014) and textural metrics derived from optical sensors in semi-arid landscapes of Mexico (St-Louis et al., 2006, 2009, 2014) and multiple habitat types in Midwestern USA (Culbert et al., 2012). All these studies identified a positive correlation between variability of habitat structure and species richness. Huang et al. (2014) have shown that spatial heterogeneity of vegetation height derived from active sensor data improves model performance over models using canopy height metrics alone. When texture statistics from optical satellite images are used, each pixel is recalculated by statistics applied to the surrounding neighborhood using a moving or fixed window algorithm. The home ranges of bird species differ considerably across body size and guilds. Image textures with multiple window sizes can help to cope with such differences in using the environment (Pearman, 2002). A suite of potentially useful statistics for calculating image textures has been proposed. Haralick (1979) developed first- and second-order statistics from the gray level co-occurrence matrix (GLCM) that explain different aspects of habitat heterogeneity. It was later shown that calculations of image textures using vegetation indices, such as the Normalized Difference Vegetation Index (NDVI), yield similar or even better results than calculations of raw bands (St-Louis et al., 2009; Wood et al., 2013). Wood et al. (2012, 2013) compared image textures derived from a high-resolution aerial photo and from a Landsat TM satellite image. They recommended using optical data derived from high-resolution sensors if habitats are structurally heterogeneous and patchily distributed but moderate-resolution sensors such as Landsat if habitat types are broader distributed. However, only few studies compared avian diversity models fitted by Lidar and optical metrics (Clawges et al., 2008; Goetz et al., 2007; Vogeler et al., 2014) or optical metrics and elevation (Sheeren et al., 2014). A comparison of complex Lidar metrics and texture metrics is important for evaluating whether cost-intensive Lidar sensors are essential to map bird diversity or whether texture metrics of optical images, which can be recurrently derived from operational satellites at lower computa-
tional cost, are able to predict bird diversity with the similar predictive power.

Most remote sensing studies use diversity proxies, such as species richness and Shannon diversity for predicting biodiversity across space (Gottschalk et al., 2005; Magurran, 1988). However, species richness alone misses much of overall diversity by neglecting the difference in functional traits of species and their roles that influence ecosystem functions (Lyashevska and Farnsworth, 2012). Since some functional traits are phylogenetically conservative (Cavender-Bares et al., 2009), phylodiversity is a first proxy for functional diversity of assemblages (Flynn et al., 2011; Lyashevska and Farnsworth, 2012). Dehling et al. (2014) showed that phylogenetic and functional diversity decreased with elevation in the tropical Andes of Peru and they suggested that this is related to factors other than climate change, i.e., to changes in habitat types or topography. Nevertheless, Shannon diversity and also simple measures of phylodiversity do not consider compositional patterns of avian communities (Fairbanks et al., 2001; Huettmann and Diamond, 2001; Marsh et al., 2010; Wiens, 1992; Wiersma and Urban, 2005). Community composition combines species richness of a given site ($\alpha$-diversity) and the change of species composition from site to site ($\beta$-diversity) (Whittaker, 1972). Various studies have shown that bird community composition provide specific information on structure, habitat requirements, and responses to local changes in habitat quality (Banks-Leite and Cintra, 2008; Cintra and Naka, 2012; Müller et al., 2009; Thiollay, 1994); e.g., Farwig et al. (2014) used the first principal component of the matrix representing the abundance of species across sites as a measure for the bird community composition. This measure was the most informative diversity proxy for modeling differences among land use-scenarios in western Kenya. As conservation resources are limited, all management decisions (e.g. area prioritization) should be based on measures of biodiversity that provide the maximum of information for the decision at issue.

Here we tested whether optical satellite data are suitable for monitoring biodiversity in areas where the availability of other remote sensing data such as full-wavelength or discrete-return Lidar and hyperspectral data are limited. We compared statistical models predicting avian diversity in a tropical mountain rainforest ecosystem at the landscape scale fitted by either Lidar metrics or optical texture metrics. In particular, we tested the predictability of three different diversity proxies: (i) Shannon diversity as a proxy of species richness and evenness, (ii) phylodiversity as a proxy for functional diversity, and (iii) community composition as a proxy for combined $\alpha$- and $\beta$-diversity.

### 2.2 Data and Methods

#### 2.2.1 Study area

The study area is located in southeastern Ecuador in the San Francisco River Valley breaching the main eastern Andean cordillera (Figure 2.1). Elevations range from 1,650 to 2,900 m a.s.l.; the climate is humid all year. Precipitation peaks from June to August; December to February are relatively drier (Beck and Kottke, 2008; Bendix et al., 2006).

The topographical complexity, heterogenic cli-
2.2 Data and Methods

Figure 2.1: Study area and sites of bird point counts in the southern Ecuadorian mountain rainforest. The main land cover classes (forest, pasture, bracken fern and barren) according to Curatola Fernández et al. (2013) are indicated.

Climate regimes along the elevational gradient, and anthropogenic disturbances have led to different forest types, namely forest in valleys and major ravines located on lower slopes (< 2,200 m a.s.l.), forests on upper slopes and along ridges (1,900–2,100 m a.s.l.), forests on ridges and upper slopes between 2,100 m and 2,250 m a.s.l., and ridge forests above 2,250 m a.s.l. up to the timberline at 2,700 m a.s.l. (Homeier et al., 2010; Paulsch and Müller-Hohenstein, 2008). The two types occurring at lower elevations are subtypes of evergreen lower montane forest, and at higher elevation forests are subtypes of evergreen upper montane forest (Homeier, 2008). These forest types differ in species richness as well as composition of plant species and therefore also of plant traits and structural characteristics. Outside the protected lands in the study area (Reserva Biológica San Francisco and Podocarpus National Park), the forest has been mostly
Table 2.1: Optical band and vegetation indices as base layers for the calculation of optical texture metrics.

<table>
<thead>
<tr>
<th>Optical metrics</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near infra-red band (NIR)</td>
<td>Pre-processed near infra-red band</td>
</tr>
<tr>
<td>Normalized difference vegetation index (NDVI)</td>
<td>(NIR - red)/(NIR + red), sensitive to chlorophyll pigments</td>
</tr>
<tr>
<td>Physiological reflectance adjusted index (PRI)</td>
<td>(red - green)/(red + green), sensitive to carotenoid/chlorophyll ratio</td>
</tr>
<tr>
<td></td>
<td>(Gamon et al., 1992; Sims and Gamon, 2002)</td>
</tr>
<tr>
<td>Anthocyanin reflectance adjusted index (ARI)</td>
<td>(red - blue)/(red + blue), sensitive to anthocyanin pigments</td>
</tr>
<tr>
<td></td>
<td>(Gitelson et al., 2001)</td>
</tr>
</tbody>
</table>

converted to pastures and bracken-infested areas (Curatola Fernández et al., 2015).

2.2.2 Avian field data and proxies for avian diversity

The diverse forest habitats of South America harbor numerous bird species. Myers et al. (2000) identified the tropical Andes as one of the world’s biodiversity hotspots, with about 1,666 bird species. Approximately 800 bird species have been observed within the Podocarpus National Park that partly overlaps with our study area (Rahbek et al., 1995).

Between 2000 and 2002, a trained observer collected data on the occurrence and abundance of birds in 30 forest sites in the study area under favorable weather conditions at 6:00–10:00 and 16:00–18:00 (UTC–5:00) (Paulsch and Müller-Hohenstein, 2008). Standardized point counts (30 min) were repeated 12 times per site and combined with mist-netting data to comprehensively assess the bird assemblages.

From this data, we calculated three different proxies of avian diversity and used them as dependent variables in all further analyses: (i) Shannon diversity index, (ii) phylodiversity, and (iii) bird community composition. We calculated phylodiversity based on 1,000 possible phylogenies (Hackett et al., 2008) of all occurring species derived from birdtree.org (Jetz et al., 2012). We extracted for each tree the phylogenetic distances between species. Subsequently, we calculated for each site and tree the mean pair-wise distance (MPD) using the package ape in R (Paradis et al., 2004). We averaged the MPD values of all possible trees, and calculated the standardized effect size (sesMPD) from the MPD values using the R package picante (Kembel et al., 2010; Webb et al., 2008). The null distribution of MPD values is based on re-shuffling the distance matrix labels across all taxa. In this study, sesMPD values serve as a measure of α-phylo-diversity. For the bird community composition, we used the first principal component representing the distribution of species along the first ordination axis (Cintra and Naka, 2012; Farwig et al., 2014). The principal component analysis was based on the covariance matrix of square-root-transformed bird counts to weight species according to their abundance (McGarigal, 2000). Results of analyses using transformed and non-transformed values did not differ in respect to the solution of the PCA (r between scores of sites on the first axis: 0.99) and therefore also in respect to the modeling exercises.
### 2.2 Data and Methods

#### Table 2.2

Texture statistics derived from gray-level co-occurrence matrices calculated from optical metrics (see Table 2.1). In the moving window approach, which shifted in all directions, two sizes were considered: $5 \times 5$ and $45 \times 45$ pixels to cope with different dependencies between species and scale of environment.

<table>
<thead>
<tr>
<th>Texture statistics</th>
<th>Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (ME)</td>
<td>$ME = \sum_{i,j=0}^{N-1} p_{i,j}$</td>
</tr>
<tr>
<td>Variance (VA)</td>
<td>$VA = \sum_{i,j=0}^{N-1} p_{i,j} (i - ME)^2$</td>
</tr>
<tr>
<td>Homogeneity (HO)</td>
<td>$HO = \sum_{i,j=0}^{N-1} \frac{p_{i,j}}{1 + (i,j)^2}$</td>
</tr>
<tr>
<td>Contrast (CO)</td>
<td>$CO = \sum_{i,j=0}^{N-1} p_{i,j} (i - j)^2$</td>
</tr>
<tr>
<td>Dissimilarity (DI)</td>
<td>$DI = \sum_{i,j=0}^{N-1} p_{i,j}</td>
</tr>
<tr>
<td>Entropy (EN)</td>
<td>$EN = \sum_{i,j=0}^{N-1} p_{i,j} (-\ln p_{i,j})$</td>
</tr>
<tr>
<td>Second moment (SM)</td>
<td>$SM = \sum_{i,j=0}^{N-1} p_{i,j}^2$</td>
</tr>
<tr>
<td>Correlation (CC)</td>
<td>$CC = \sum_{i,j=0}^{N-1} p_{i,j} \left( \frac{(i-ME)(j-ME)}{\sqrt{VA_i VA_j}} \right)$</td>
</tr>
</tbody>
</table>

*With $p_{i,j} = V_{i,j} \sum_{i,j=0}^{N-1} V_{i,j}$, where $V_{i,j}$ is the value in cell $i$, $j$ and $N$ is the number of rows or columns.

#### 2.2.3 Remote sensing indicators

To fit the models, we calculated two predictor datasets: (a) texture metrics of optical images derived from a satellite-borne passive sensor and (b) Lidar metrics derived from an airborne active sensor. The texture metrics were based on a Quickbird scene acquired on 22 October 2010 at 9:54 local time under clear weather conditions from Digital Globe. Images were preprocessed, including geometric, atmospheric, and topographic corrections (Curatola Fernández et al., 2015).

In spite of the time lag between field data and image data acquisition, forest habitats in our study area as well as neighboring habitats have not changed during the time between sampling of bird assemblages and remote sensing data (Appendix A Figure A.1). Owing to computation time, we decided to use only the corrected near-infrared (NIR) channel and three vegetation indices to calculate image textures (Table 1), namely (i) the NDVI to account for vegetation productivity and health, (ii) an approximation of the physiological reflectance index (PRI) to account for photosynthetic efficiency (Gamon et al., 1992) and carotenoid/chlorophyll ratio (Sims and Gamon, 2002), and (iii) an approximation of the anthocyanin reflectance index (ARI) to account for anthocyanin pigments (Gitelson et al., 2001; Sims and Gamon, 2002). Based on the four mentioned base layers, we first calculated texture statistics derived from the gray level co-occurrence matrix using the glcm package in R (Zvoleff, 2015). The calculated
statistics encompassed mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation (Haralick, 1979) for four different sizes of the moving window to match different scales of bird habitats among feeding guilds (Table 2.2). Correlation is one of the most independent algorithms among all texture statistics and sometimes returns no-data values since the software cannot handle 0 (no correlation) as denominator in the correlation equation. For those metrics with sparse undefined areas, we applied a filter function in the \textit{raster} package in R to recalculate no-data pixels by the mean of their neighboring pixels. However, we excluded those metrics with a large number of undefined correlations. In the second step, we reduced the window sizes to two numbers (smallest and biggest); the model accuracies did not change. Thus, considering the spatial resolution of Quickbird images of 2.5 m per pixel (after transformation for topographic correction), we chose $5 \times 5$ and $45 \times 45$ pixel moving windows to match a surrounding of $12.5 \times 12.5$ m$^2$ and $112.5 \times 112.5$ m$^2$, respectively.

The second dataset is based on vertical forest structure metrics derived from Lidar data (Silva and Bendix, 2013). Airborne laser scanning (ALS) data were collected using a Leica ALS-50-II laser scanner during two campaigns in March and November 2012 with a Eurocopter AS350B2 Ecureuil helicopter. The point cloud density counts were at least 10 pulses per 1 m$^2$, considering steep slopes and valleys, and was classified prior to metrics calculation, so that only vegetation (not ground) points were considered (Silva and Bendix, 2013). From the point cloud with x, y, z values, derived from Lidar pulse returns, we calculated a suite of measures related to vertical vegetation structure using the FUSION software (McGaughey, 2009). In our analyses, we chose metrics according to their success in published bird diversity and forest ecology studies (Table 2.3; Goetz et al., 2007; McGaughey, 2009; Vogeler et al., 2014). Additionally, we included the intensity of back reflected energy, which was also recorded by the laser scanner. Since some metrics had a few cells with no-data, we replaced these cells with computed focal statistics of type mean. We considered texture in the Lidar metrics to take into account the spatial arrangement of vertical structure and to increase comparability between Lidar metrics and optical texture metrics. We measured the texture statistics mean and variance of all Lidar metrics within two moving window sizes of $13 \times 13$ and $113 \times 113$ pixels.

### 2.2.4 Analyses

Approaches considering image texture have to deal with high dimensionality and intercorrelations which decrease model accuracy in statistical models using standard methods (Beyer et al., 1999). Partial least-squares (PLS) regression is a machine-learning algorithm with a prior feature space transformation that minimizes a least-squares cost function. Thus, PLS regression is able to cope with a higher number of predictor variables than predictants and strong co-linearity among predictors (Andersen and Bro, 2010; Carrascal et al., 2009; Mehmood et al., 2012). We used the automatic PLS regression approach implemented in the R package \textit{autopls} (Schmidtlein et al., 2012). The algorithm
Table 2.3: Vertical structure metrics for fitting Lidar models. For each index, first-order statistics, mean, and variance were calculated using a $13 \times 13$ and a $113 \times 113$ pixel moving window (see Table 2.2).

<table>
<thead>
<tr>
<th>Lidar metrics</th>
<th>Abbr.</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>DEM</td>
<td>Digital elevation model</td>
</tr>
<tr>
<td>Slope</td>
<td>SLOPE</td>
<td>Gradient of DEM in degree</td>
</tr>
<tr>
<td>Canopy height</td>
<td>CH</td>
<td>Maximum height</td>
</tr>
<tr>
<td>Median height</td>
<td>MH</td>
<td>The median canopy height</td>
</tr>
<tr>
<td>Vertical distribution ratio</td>
<td>VDR</td>
<td>$(CH-MH)/CH$</td>
</tr>
<tr>
<td>Density metrics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E57</td>
<td></td>
<td>Percentage of all returns above mode</td>
</tr>
<tr>
<td>E50</td>
<td></td>
<td>Percentage of returns above 3 m height</td>
</tr>
<tr>
<td>E55</td>
<td></td>
<td>Percentage of first returns above mode</td>
</tr>
<tr>
<td>Canopy closure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E58</td>
<td></td>
<td>$(All\ returns\ above\ mean)/(Total\ first\ returns) \times 100$</td>
</tr>
<tr>
<td>E59</td>
<td></td>
<td>$(All\ returns\ above\ mode)/(Total\ first\ returns) \times 100$</td>
</tr>
<tr>
<td>E51</td>
<td></td>
<td>$(All\ returns\ above\ 3\ m)/(Total\ first\ returns) \times 100$</td>
</tr>
<tr>
<td>Canopy relief ratio</td>
<td>E68</td>
<td>$(mean - min) / (max - min)$</td>
</tr>
<tr>
<td>Intensity</td>
<td>INT</td>
<td>Mean intensity of backscattered laser pulse</td>
</tr>
</tbody>
</table>

was trained with all 30 sites using a leave-one-out (LOO) validation, where one observation is omitted at a time in the models to provide prediction errors. As a wrapper function, the implemented variable selection was used to improve model performance and interpretability, and to reduce computation time (Andersen and Bro, 2010). We applied an optimization procedure using a filtering based on thresholds for backward selection, i.e., significance, variable importance in the projection (VIP), or both (Chong and Jun, 2005; Mehmood et al., 2012). Each of the three diversity proxies was regressed against the two remote sensing datasets using PLS. To compare the resulting six models, we used the LOO $R^2$ values and LOO root-mean-square errors (RMSE) as comparative measures for predictive power and prediction error. We compared significance and magnitude among predictors in our top models utilizing the weighted regression coefficients based on significance considerations from jackknifing (Mehmood et al., 2012). In PLS regression, the importance of predictors is proportional to its distance from the origin in the loading space. In the final step, we tested for spatial autocorrelation and applied Moran’s I of the residuals of each bird diversity model using the R package `spdep` (Bivand and Piras, 2015). We detected no significant spatial autocorrelation even after considering various neighborhood combinations.

For models with a high predictive power, we predicted spatial patterns of the diversity proxy within the defined area of interest (Figure 2.1). To do so, the fitted model was applied to a raster stack that included all predictors defined by the backward selection process using the predict function in the R package `raster` (Hijmans et al., 2015). For the prediction maps, we masked out non-forested areas because bird abundance was only sampled in forest sites. The resulting raster thus contains pixels with predicted values of the diversity proxies derived from the fitted model.
2.3 Results

2.3.1 Diversity proxies

A total of 147 different bird species were recorded within the 30 study sites across an elevation gradient of 700 m. The three diversity proxies measured showed high variation within study sites (Table 2.4). While phylodiversity was independent from the other proxies, the bird community, however, was significantly related to Shannon diversity ($r = 0.58$, $p < 0.001$), showing that the bird community is not only a proxy for $\beta$-diversity, but also for $\alpha$-diversity (Appendix A Figure A.2). The first principal component which was based on abundance data explained more than 20% of the overall variance in species abundances across all sites (Appendix A Table A.1). *Cyanocorax yncas* (Green Jay), *Myioborus miniatus* (Slate-throated Whitestart), *Henicorhina leucophrys* (Grey-breasted Wood-wren), *Basileuterus tristriatus* (Three-stribed Warbler) showed positive loadings in the first principal component whereas *Chlorospingus ophthalmicus* (Common Bush-tanager), *Scytalopus unicolor* (Unicoloured Tapaculo), *Synallaxis unirufa* (Rufous Spinetail), *Pionus senilis* (Crowned Parrot) showed negative loadings (Appendix A Table A.2). We searched for a pattern among morphological and life history traits, such as species size/weight, foraging strata, or center of abundance, but the traits of neotropical birds collected in Stotz et al. (1996) and Dunning (2008) did not explain the scores along the first axis (Appendix A Table A.3).

2.3.2 PLS regression model comparison

A suite of 52 Lidar metrics and 63 optical texture metrics were generated to test and compare the predictability of bird diversity. We found different patterns of sensor suitability for the three selected diversity proxies, ranging from highest LOO $R^2$ values for community composition to lowest values for phylodiversity (Table 2.5). With both datasets community composi-

### Table 2.4: Summary statistics for proxies of avian diversity. Mean (CI) = mean with the 95% confidence intervals.

<table>
<thead>
<tr>
<th>Diversity measure</th>
<th>Median</th>
<th>Mean (CI)</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon diversity</td>
<td>2.99</td>
<td>3.04 (2.85 - 3.27)</td>
<td>2.53</td>
<td>3.73</td>
</tr>
<tr>
<td>Phylo-diversity</td>
<td>-0.15</td>
<td>0.15 (-0.49 - 1.01)</td>
<td>-1.92</td>
<td>2.27</td>
</tr>
<tr>
<td>Bird community</td>
<td>0.36</td>
<td>0 (-1.53 - 1.97)</td>
<td>-3.70</td>
<td>3.89</td>
</tr>
</tbody>
</table>

### Table 2.5: Comparison of partial least-square regression models of the three diversity proxies using Lidar and texture statistics from optical satellite images. The best model results are in bold.

<table>
<thead>
<tr>
<th>Diversity measure</th>
<th>Dataset</th>
<th>No. of predictors</th>
<th>No. of latent vectors</th>
<th>$R^2$</th>
<th>RMSE</th>
<th>LOO $R^2$</th>
<th>LOO RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon diversity</td>
<td>Lidar only</td>
<td>5</td>
<td>2</td>
<td>0.26</td>
<td>0.85</td>
<td>0.07</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Texture only</td>
<td>7</td>
<td>4</td>
<td>0.57</td>
<td>0.64</td>
<td><strong>0.42</strong></td>
<td>0.75</td>
</tr>
<tr>
<td>Phylo-diversity</td>
<td>Lidar only</td>
<td>13</td>
<td>2</td>
<td>0.60</td>
<td>0.62</td>
<td><strong>0.35</strong></td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Texture only</td>
<td>8</td>
<td>2</td>
<td>0.24</td>
<td>0.86</td>
<td>0.09</td>
<td>0.94</td>
</tr>
<tr>
<td>Bird community</td>
<td>Lidar only</td>
<td>11</td>
<td>3</td>
<td>0.86</td>
<td>0.37</td>
<td><strong>0.81</strong></td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Texture only</td>
<td>8</td>
<td>2</td>
<td>0.85</td>
<td>0.38</td>
<td><strong>0.80</strong></td>
<td>0.44</td>
</tr>
</tbody>
</table>
2.3 Results

Figure 2.2: Regression coefficient plots of the top (best) models identifying predictor variables and their significance in the corresponding model. With * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$

The backward selection process identified different combinations of predictors for each diversity proxy (Appendix A Table A.4). For Lidar models, no patterns of selected window sizes were found, but as a result of co-linearity, often both window sizes were selected together within the latent vectors. Likewise, no general pattern was found in the selection of specific...
Figure 2.3: Prediction maps of the bird community by assessing (a) Lidar metrics and (b) optical texture metrics. Colors are scaled according to the lowest minimum value and highest maximum value of the two predictions. (c) Elevation (DEM) was one of the main predictors of the Lidar model. (d) The difference between the Lidar model and optical texture models. Higher values indicate that the bird community was rated higher by the Lidar model, and lower values indicate that the bird community was rated higher by optical texture models.

Texture statistics among the metrics, but a high interchangeability of most texture metrics was found between latent vectors owing to high colinearity among texture statistics. Texture metrics derived from the larger window size were more frequently selected as predictors.

Weighted regression coefficients of the top models identified magnitude and sign of the relationship between predictors and observed diversity (Figure 2.2). The Shannon model fitted by texture metrics included seven predictors containing homogeneity as well as heterogeneity textures. All predictors in this model were generated using the $45 \times 45$ pixel window size. Most significant ($p < 0.05$) was the contrast (heterogeneity) of NIR, which was positively related to Shannon diversity, followed by correlation, which calculates linear dependencies of PRI ($p < 0.1$) and is the most independent texture measure. Negatively related were second
moment (homogeneity) of PRI and entropy (heterogeneity) of NIR. In the community model, heterogeneity textures derived from NIR and correlation of PRI were highly associated with the bird community. The most significant predictors in Lidar community models were elevation and the median canopy height (MH) (both \( p < 0.001 \)), followed by canopy relief ratio (E68) and canopy closure (E58) (both \( p < 0.1 \)). For phylodiversity, 13 Lidar metrics were included in the model. In contrast to the community model, slope had a negative effect on phylodiversity. In addition, the effects of density metrics (E55 and E57) depended on the window size of the texture approach, where a larger size was related to a negative association, and a smaller size was related to a positive association.

### 2.3.4 Prediction maps bird community

Because of the high predictability of the community composition, we mapped the two bird community models across our study area (Figure 2.3a,b). Since the two models differed in their predicted range (Table 2.6), we normalized the color to their minimum and their maximum range to make them quantitatively comparable. The Lidar model predicted lower values of the bird community than the optical texture model, particularly in the northwestern and southeastern parts of the study area, where forests are at higher elevations and have a lower vertical structure. The optical texture model predicted higher bird community values, particularly at lower elevations near the San Francisco River and on higher slopes and in ravine forests next to pastures. We subtracted the Lidar model from the optical model to identify major differences between the two models (Table 2.6, Figure 2.3). Highest differences were found in the valleys of the northwestern region, next to non-forested (mostly pastures) patches or higher elevations above 2,300 m a.s.l.

### 2.4 Discussion

Our analysis of three different proxies for bird diversity in a tropical mountain forest ecosystem in southeastern Ecuador revealed clear differences in the suitability of sensor information for predicting the various aspects of biodiversity. The predictive power of Lidar and optical texture models particularly differed when the proxies Shannon diversity and phylodiversity were used. Bird community composition was best predicted with equal predictive power irrespective of the independent dataset used. In the following we address the relation between remotely sensed variables and forest structure to further explain the spatial patterns of bird diversity in our study area.

#### 2.4.1 Forest structure and spectral respond

In contrast to previous bird diversity studies assessing remote sensing data in deserts or even boreal forests, our study area consisted only of forest habitats which are vertically more diverse and complex. Our study area is characterized by two gradients of forest types. First, the elevational gradient where lower tree density is found with higher elevation and ridges (Homeier et al., 2010). Forests above 2,250 m a.s.l. are characterized by the dominance of a single tree species (Purdiaea nutans Planch) due to harsher
environmental conditions such as wind speed (Homeier et al., 2010; Wagemann et al., 2015). The elevational change of forest structure associates lower chlorophyll content at higher elevations which in return implies less reflectance of NIR and NDVI values. In the meantime, solar ultraviolet radiation increases with altitude (Piazena, 1996) which leads to more carotenoid pigments within vegetation and in turn results in higher PRI values (Gamon et al., 1992; Gitelson et al., 2001). The second gradient relates to topography. Ravine and ridge forest differ in their edaphic conditions owing to soil nutrient dispersion in ridge forests and concentration in ravine forests (Werner and Homeier, 2015). Thus, we expect low reflection of NIR due to low chlorophyll content in ridge forests with lower tree density and tree height. Clearly, elevation, and topography can affect forest structure as well as the spectral response at the satellite.

### 2.4.2 Shannon diversity

With respect to the most significant optical predictors, i.e., correlation of PRI and contrast of NIR, we deduce an association of high Shannon diversity and forests with different tree layers and a heterogeneous canopy structure. We conclude that a higher number of species is associated with lower montane forests in valleys and major ravines that account for a higher structural variety that responds to variations in the NIR. As noted by Culbert et al. (2012), models fitted by image textures in landscapes with low vertical variability gain higher predictability than in vertically complex landscapes. In our study, the predictive power of modeling Shannon diversity was similar to predictions of overall species richness in earlier studies of desert landscapes or multiple habitats (Culbert et al., 2012; St-Louis et al., 2009).

The results obtained with our Lidar models did not support studies in which topographical metrics successfully predicted avian species richness (Goetz et al., 2007; Huang et al., 2014). In the study of Vogeler et al. (2014), a higher number of returns yielded a higher predictive power. Lesak et al. (2011) used low density discrete-return (1–3 returns) Lidar data, which explained approximately 18% of variance (adj. R²) in song bird species richness in deciduous forests of southern Wisconsin, USA. Although our Lidar data had up to 10 returns per m², Lidar metrics explained only 9% of Shannon diversity. This is probably related to the relatively low return density of Lidar pulses in the understory layers (Müller et al., 2010). High closure in the upper canopy reduces return densities in the lower tree layers, and thus, understory characteristics that are important for specific feeding guilds are neglected.

### 2.4.3 Phylodiversity

The negative effect of topography on phylodiversity observed in our study supports the results of Dehling et al. (2014), in which higher values of phylodiversity were found at lower elevations and lower slopes. In our study, Lidar predictors combined topographic aspects and vertical density of vegetation. We found that higher density in the upper strata had a positive effect on the distribution of genetic diversity and suggest that more diverse micro-habitats can harbor species from different feeding guilds in the canopy and the understory. Texture metrics of
optical images are reduced to spatial variations in reflectance and cannot cope with structural forest components in different strata to explain micro-habitat structure, and thus were not able to predict phylodiversity in our study area.

2.4.4 Bird community

Our community models indicated that the analyzed species composition was more strongly associated with remotely sensed habitat structure than Shannon diversity and phylodiversity, which supports the results of Banks-Leite and Cintra (2008). In their study, bird community measures of frugivorous and nectarivorous birds assign lower importance to evenly distributed species along the study area and the observed bird assemblage was composed of only a few species associated with canopy openness and elevation. In our study, the bird community was mainly composed of four species highly contributing to the score of the first component (Appendix A Table A.2). Since the environmental gradient of our study area is fine-scaled resulting in many micro-habitats, it is likely that overall species traits are not highly resolved enough to explain species composition at a landscape scale.

Owing to the correlation of Shannon diversity and bird community, optical predictors of the Shannon diversity models and the community models partly correspond to each other resulting in a spatial overlap of high Shannon diversity and bird community. Texture metrics of optical images described the habitat of the bird community with higher intensity and heterogeneity of NIR and higher homogeneity of PRI which is present in the lower montane forests in valleys and ravines and partly along ridges and upper slopes.

According to the Lidar model, forest habitats of the bird community are located at lower elevations and higher mean canopy height. Low community composition, mainly composed by highly negatively loaded species (Appendix A Table A.3), is associated with high altitudes, low median canopy height and high vertical variations. Canopy closure metrics (E58 and E51) define that the tree crowns are not closed due to scattered large trees and a diverse upper vegetation strata.

The prediction maps differed in the predicted bird community values from the observed ranging from −3.8 to 3.7. Lidar models overestimated the maximum community value with a minimum value of −1.98, and texture metrics of optical images underestimated the community with a maximum value of 2.18. The differences in the predicted range of the community between Lidar and optical texture models are associated with higher elevations above 2,300 m a.s.l. The only two observations were located around 2,500 m a.s.l.; the lack of observations above this eleva-

### Table 2.6: Summary statistics of predicted bird community models fitted by Lidar and optical texture metrics and the difference between the two calculated values.

<table>
<thead>
<tr>
<th>Bird community maps</th>
<th>SD</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Lidar model</td>
<td>0.85</td>
<td>0.41</td>
<td>-3.5</td>
<td>1.98</td>
</tr>
<tr>
<td>(b) Optical texture model</td>
<td>0.74</td>
<td>0.12</td>
<td>-2.18</td>
<td>2.5</td>
</tr>
<tr>
<td>(d) Difference map</td>
<td>0.61</td>
<td>0.30</td>
<td>-2.50</td>
<td>2.15</td>
</tr>
</tbody>
</table>
tion makes predictions difficult. Likewise, community values for the fragmented forest close to pastures predicted by the two models also greatly differ. Texture metrics of optical images predicted significantly lower community values than Lidar metrics. A low number of observations in these areas again result in uncertainties.

2.4.5 Limitations and recommendations

Given the low number of sampling sites, we detected an over-fitting of most models when we compared the calculated with the validated coefficient of determination in the response plots (Appendix A Figure A.3). For an adequate validation, in particular in those areas where our two models for the bird community differed considerably, further bird surveys are needed. This is particularly the case in forests above 2,300 m a.s.l. and in forests close to pastures and bracken fern. Bird diversity may be affected by disturbances, such as cattle grazing and presence of bracken fern (Maya-Elizarrarás and Schondube, 2015), which can affect forest habitat structure and in turn its spectral response. Further investigations of fragmented forest sites and the upper montane forests might improve bird diversity models.

Owing to the scarce availability of remote sensing data in our study area, we did not test full-wavelength Lidar data, which might perform better in describing small-scale habitat patterns (Goetz et al., 2007), and hyperspectral data, which might perform better in characterizing the habitat composition defined by tree species (Carlson et al., 2007). Compared to these costly sensor sources, optical satellite or airborne images are more accessible and more effective over a longer period and for large-scale monitoring (Nagendra and Rocchini, 2008; Turner et al., 2003). Where high-resolution images are not available, satellite data at medium resolution such as Landsat have been successfully used to model forest habitat structure and disturbance related to avian habitat using either time-series of cloud-cleared image mosaics (Helmer et al., 2010) or image textures (Culbert et al., 2012; Wood et al., 2012, 2013). Wall-to-wall coverage, spectral and temporal resolution, in particular in combination with a compositing or time-series approach, makes Landsat a valuable tool in ecological driven studies (Cohen and Goward, 2004; Irons et al., 2012; Wulder et al., 2008). However, in landscapes such as in the present study area where forest habitat types change rapidly and show structural complexity, the spatial resolution of 30 m per pixel might not cover small-scale changes in habitat structure and thus avian diversity (Wood et al., 2012, 2013). Since optical texture models performed equally or even better in modeling certain diversity proxies than Lidar models, we strongly suggest the use of optical satellite or airborne high-resolution data in combination with texture statistics to model bird diversity proxies such as Shannon diversity or community compositions at a landscape scale.

As an improvement to the use of texture metrics, we recommend using uncorrelated texture statistics to save computation time (St-Louis et al., 2006), and in a similar manner, reducing window sizes to a single size. Variograms can help to identify the maximum window size (Kuemmerle et al., 2009), or one can choose an ecologically meaningful size that character-
izes the potential maximum habitat of observed species.

2.4.6 Conclusions and conservation implications

The relationship between avian biodiversity and metrics derived from remote sensing depends on both the sensor type and diversity proxy. The good performance of texture metrics observed in our study underpins the use of optical remote sensing data as a substitute for Lidar data in habitat diversity studies, thereby reducing both costs and pre-processing time. This is beneficial particularly for areas where other sensor data, e.g., Lidar or hyperspectral data, are lacking. However, optical texture models were limited to certain diversity measures, such as Shannon diversity and community composition. In addition, the spectral respond of the forest was related to structural characteristics which significantly correspond to the landscape of our study area. Models of the bird community composition revealed a high potential in identifying relationships between some specific species assemblages that are not found for overall species or genetic diversity. Relationships between the environment and present avian species strongly depend on the identity of their composing communities (Filloy et al., 2010). Hence, the integration of avian community composition in addition to proxies for species richness in conservation planning could enhance the understanding of bird assemblages in terms of habitat preferences and response to habitat changes (Banks-Leite and Cintra, 2008; Cintra and Naka, 2012; Farwig et al., 2014). This is particularly of importance at a landscape scale, where a high number of species is evenly distributed and only a few of the total species depend on micro-habitat forest structure. Compared to spatial reflectance variations, Lidar metrics point out the importance of vertical structural complexity related to ecological traits. Up-coming airborne and spaceborne sensor missions with passive and active sensors on board will strengthen biodiversity-driven research (O’Connor et al., 2015), in particular genetic diversity related to functional traits. Nevertheless, field data such as a long-time monitoring of bird abundance is necessary to fully understand spatial patterns of bird diversity in the mountain rainforest of southeastern Ecuador.

Acknowledgements

This study was conducted in the framework of the Research Unit 823-825 Platform for Biodiversity and Ecosystem Monitoring and Research in South Ecuador of the German Research Foundation (DFG), subproject C2 (BE1780/34-1, BR1293/11, FA925/7-1, ZI698/8-1). We are grateful to the DFG for funding our project, to the Ecuadorian Ministry of the Environment (MAE) for permission to conduct research, and to the foundation Nature and Culture International (NCI) for logistic support. We acknowledge the comments of three anonymous reviewers that helped to improve the manuscript.

References


Irons, J.R., Dwyer, J.L., Barsi, J.A., 2012. The next Landsat satellite: The Landsat Data Continu-


Popp, A., Humpenöder, F., Weindl, I., Bodirsky, B.L., Bonsch, M., Lotze-Campen, H., Müller, C., Biewald, A., Rolinski, S., Stevanovic, M.,


Silva, B., Bendix, J., 2013. Remote sensing of vegetation in a tropical mountain ecosys-
Chapter 2 Predicting bird diversity by textural metrics in a tropical mountain ecosystem


Zvoleff, A., 2015. Glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R.
Chapter 3

Remote sensing improves prediction of tropical montane species diversity but performance differs among taxa

Christine I. B. Wallis, Gunnar Brehm, David A. Donoso, Konrad Fiedler, Jürgen Homeier, Detlev Paulsch, Dirk Süssenbach, Yvonne Tiede, Roland Brandl, Nina Farwig, Jörg Bendix

published in Ecological Indicators (2017), 83, 1, 538–549
(doi: 10.1016/j.ecolind.2017.01.022)
Abstract

Texture information from passive remote sensing images provides surrogates for habitat structure, which is relevant for modeling biodiversity across space and time and for developing effective ecological indicators. However, the applicability of this information might differ among taxa and diversity measures. We compared the ability of indicators developed from texture analysis of remotely sensed images to predict species richness and species turnover of six taxa (trees, pyraloid moths, geometrid moths, arctiinae moths, ants, and birds) in a megadiverse Andean mountain rainforest ecosystem. Partial least-squares regression models were fitted using 12 predictors that characterize the habitat and included three topographical metrics derived from a high-resolution digital elevation model and nine texture metrics derived from very high-resolution multi-spectral orthophotos. We calculated image textures derived from mean, correlation, and entropy statistics within a relatively broad moving window (102 m × 102 m) of the near-infrared band and two vegetation indices. The model performances of species richness were taxon dependent, with the lowest predictive power for arctiinae moths (4%) and the highest for ants (78%). Topographical metrics sufficiently modeled species richness of pyraloid moths and ants, while models for species richness of trees, geometrid moths, and birds benefited from texture metrics. When more complexity was added to the model such as additional texture statistics calculated from a smaller moving window (18 m × 18 m), the predictive power for trees and birds increased significantly from 12% to 22% and 13% to 27%, respectively. Gradients of species turnover, assessed by non-metric two-dimensional scaling (NMDS) of Bray-Curtis dissimilarities, allowed the construction of models with far higher predictability than species richness across all taxonomic groups, with predictability for the first response variable of species turnover ranging from 64% (birds) to 98% (trees) of the explained change in species composition, and predictability for the second response variable of species turnover ranging from 33% (trees) to 74% (pyraloid moths). The two NMDS axes effectively separated compositional change along the elevational gradient, explained by a combination of elevation and texture metrics, from more subtle, local changes in habitat structure surrogated by varying combinations of texture metrics. The application of indicators arising from texture analysis of remote sensing images differed among taxa and diversity measures. However, these habitat indicators improved predictions of species diversity measures of most taxa, and therefore, we highly recommend their use in biodiversity research.

3.1 Introduction

Information derived from remote sensing (RS) provides cost-effective proxies for primary productivity and habitat structure (Rocchini et al., 2016, 2015; Wang et al., 2010). Species occurrence of individual species and species diversity are often correlated to these proxies (Cintra and Naka, 2012; Couteron et al., 2005; Estes et al., 2010; Goetz et al., 2007; Mairota et al., 2015; Rocchini et al., 2010; Tews et al., 2004). Therefore, RS provides useful information for models of ecological variables across large extents with a high spatial resolution. Such spatially-explicit models are of considerable importance in conservation planning if recurrent RS information is available as they provide maps and offer effective indicator systems for area-wide monitoring. However, the success of these models varies considerably among taxa and modeled variables of biodiversity. A deeper understanding of this variation in predictability of diversity measures would be helpful for planning and es-
3.1 Introduction

 Establishing monitoring systems for documenting environmental change, especially in biota where biodiversity inventories are difficult to achieve.

 Particularly the use of RS texture metrics has strengthened statistical models of biodiversity (Culbert et al., 2012; Estes et al., 2010; Wallis, 2016; Wood et al., 2013). In textural approaches, a new value is assigned to each pixel and characterizes the distribution of spectral values in a particular neighborhood, which is defined by a moving or fixed window (Haralick, 1979). Depending on the considered textural feature, which ranges from simple metrics (e.g., mean, variance) to complex metrics (e.g., contrast, correlation), such variables characterize different spatial aspects of habitat structure (e.g., habitat heterogeneity). For example, image textures based on very high-resolution optical imagery successfully predict and map the structure of forests (Wood et al., 2012) and distributional patterns of bird diversity (St-Louis et al., 2014). Models of a montane forest in southwestern Colorado that include texture metrics from RS are more strongly correlated with biomass than models using topographical or spectral metrics (Kelsey and Neff, 2014). Similar results have been obtained for mature biomass in a moist tropical forest (Lu, 2005). Therefore, textural information from RS images might address the relationship between environment and biodiversity more effectively than raw spectral bands or common vegetation indices.

 Tropical mountain rainforests, particularly Andean rainforests, are among the most diverse and threatened biodiversity hotspots of the world (e.g., Brehm et al., 2016; Tapia-Armijos et al., 2015). Studies of similarly diverse systems have investigated elevation and topography as predictors of biodiversity, and have successfully modeled the occurrence of certain tree species and species richness of moths and ants (e.g., Kübler et al., 2016; Malsch et al., 2008; Nakamura et al., 2015). However, the results are highly taxon dependent, and some taxa are difficult to predict from simple environmental variables (Fiedler et al., 2008; Tiede et al., 2016a). Thus, models of tropical diversity would benefit from the inclusion of structural habitat information.

 A great challenge in tropical diverse systems is the assembly of meaningful biodiversity data. The extraordinary high species richness and the low availability of taxonomic and ecological information for most of the species forces ecological studies in tropical rainforest ecosystems to target well-known taxa, e.g., woody plants (Homeier et al., 2010), or taxa such as ants or birds that occupy different trophic levels within the food webs (Gerlach et al., 2013; Kati et al., 2004; Schuld et al., 2014; Sekercioğlu et al., 2016; Tiede et al., 2017; Donoso and Ramón, 2009).

 Most RS-based diversity research has focused on measures of alpha-diversity and has ignored community structure. Changes in species composition along environmental gradients are measured by a variety of metrics of species turnover, ranging from dissimilarity measures to scores along ordination axes (Socolar et al., 2016; Whittaker, 1972; Brehm and Fiedler, 2004). Various studies have shown that the composition of species usually provides detailed information on habitat characteristics (Banks-Leite and Cintra, 2008; Cintra and Naka, 2012; Farwig et al., 2014; Müller et al., 2009; Thiollay, 1994). RS-based habitat indicators might improve predictions of
species turnover as spectral distances are, for example, strongly correlated to patterns of floristic species composition among sites in Ecuadorian Amazonia (Tuomisto et al., 2003). Unfortunately, the computation of spectral distances and species similarities among sites and the mapping of distance-based species turnover is time consuming because the number of spectral distances increases with the square of the number of sampled sites (Rocchini et al., 2016). A number of studies, therefore, have assessed species turnover using ordination techniques (Farwig et al., 2014; Feilhauer and Schmidtlein, 2009; Gu et al., 2015; Muenchow et al., 2013; Wallis, 2016). For instance, Feilhauer and Schmidtlein (2009) performed a detrended correspondence analysis to identify different gradients in the composition of vegetation. Scores of sites along the ordination axes that represent these environmental gradients were regressed against topographical and spectral metrics calculated for the sites. Ordination techniques, therefore, might be the superior choice to assess species turnover when RS information are used to produce continuous maps of environmental gradients, e.g., the compositional change of bird species along habitat structure (Wallis, 2016).

Here, we investigated models that consider species richness and ordinations of compositional change using non-metric multidimensional scaling (NMDS) as a measure of species turnover of trees, moths (Pyraloidea, Geometridae, Arctiinae), ants, and birds in a tropical mountain rainforest ecosystem. We fitted partial least-squares regressions for all taxa and diversity measures separately by assessing topographical metrics derived from a digital elevation model and image texture metrics derived from an airborne multi-spectral sensor. Our aim was to compare the predictability of species richness and species turnover across the six taxa. We identified differences among the selected taxonomic groups and examined which combination of habitat indicators served well for each taxonomic group as well as for the selected diversity measure. Our findings provide guidelines for the development of a RS-based indicator system for monitoring biodiversity in response to environmental changes in complex tropical forests.

3.2 Data and Methods

3.2.1 Study area

The study area is located in southeastern Ecuador, an area known for its high climatic and environmental heterogeneity and high levels of species richness with numerous endemic taxa (Bendix and Beck, 2016; Brehm et al., 2016, 2008; Tapia-Armijos et al., 2015; Werner and Homeier, 2015, Figure 3.1). The climate is humid throughout the year, with mean annual precipitation between 2,200 mm at elevations of ~1,000 m a.s.l. and 4,500 mm at ~3,000 m a.s.l. Mean annual air temperature decreases from 20° C at ~1,000 m a.s.l., to 15.5° C at ~2,000 m a.s.l., to 9.5° C at ~3,000 m a.s.l. (Beck et al., 2008; Bendix et al., 2008). Precipitation peaks from June to August, and the period from December to February is relatively dry (Beck and Kottke, 2008; Bendix et al., 2006).

Topographical complexity, heterogeneity of climate regimes along the elevational gradient, and anthropogenic disturbances have formed various forest types: evergreen premontane
3.2 Data and Methods

Figure 3.1: A) Study sites and B) plots of all taxa sampled, covering different elevation levels and environmental gradients. A) Color images represent the true color composite of SIGTIERRAS orthophotos; gray areas are non-vegetation areas or mask out pixels where the sensors signal was low owing to casted shadows. B) The number of samples (n) stratified along the elevational gradient for each taxon.

forest (≤1,300 m a.s.l., Bombuscaro site), evergreen lower montane forest (between 1,300 and 2,100 m a.s.l., San Francisco site), and evergreen upper montane forest (> 2,100 m a.s.l., San Francisco and Cajanuma sites), with characteristic subtypes along valleys, ravines, ridges, and anthropogenic replacement systems (Homeier et al., 2008). These forest types differ in species richness, floristic composition, and structural characteristics, e.g., canopy height decreases with elevation. We focused on taxonomic groups sampled on different forest plots and elevations in the Reserva Biológica San Francisco, or additionally at Bombuscaro and Cajanuma which are parts of the Podocarpus National Park (Figure 3.1; Table 3.1). Outside the protected lands in the study area, the forest has been mostly converted during the last few decades to pastures or further degraded to bracken-infested areas (Curatola Fernández et al., 2015; Tapia-Armijos et al., 2015).
3.2.2 Sampling of taxa

All woody plants (including palms and arborescent ferns) with a diameter at breast height (dbh) $\geq 5$ cm and a height $\geq 1.3$ m were sampled on 50 permanent plots each covering $20 \times 20$ m between 2007 and 2008 (Wallis and Homeier, 2017). GPS points were taken in the middle of each plot. All plots were located in mature forest without visible human disturbance. Tree individuals were identified to the species or morphospecies level when possible.

Nocturnal moths were sampled on 20 plots during three periods between 1999 and 2000 using light traps equipped with two backlight 15 W tubes operated between 18:30 and 21:30 local time (UTC-5:00; Brehm, 2002; Brehm and Fiedler, 2010; Suessenbach, 2003; Suessenbach and Fiedler, 2010a,b). Catches were restricted to periods from three days after full moon until five days before full moon. Specimens were sorted to morphospecies within three diverse taxa: geometrid moths ($\text{Geometridae}$), pyraloid moths ($\text{Pyraloidea}$), and arctiinae moths ($\text{Erebidae: Arctiinae}$). Previous analyses of the same samples have shown that these taxa differ in diversity patterns, body size, host plant affiliations, and other traits (Fiedler et al., 2008). To account for such differences, we treated these taxa separately to compare the predictability of biodiversity patterns among putatively rather homogeneous groups of insects.

Ants ($\text{Formicidae}$) were sampled on 27 plots in two seasons in 2014 using a rapid assessment method with nutrient baits following Peters et al. (2014) and described in detail in Tiede et al. (2017) and Wallis (2016). In short, we exposed six different nutrient baits [$\text{H}_2\text{O}$, $\text{NaCl}$, CHO (sucrose), protein (glutamine), CHO-protein mix, and lipids (olive oil)] in the wet and the dry seasons. The ant fauna of Ecuador consists of species that are taxonomically poorly defined (Salazar et al., 2015). Therefore, we built a COI (mitochondrial cytochrome c oxidase I) barcode reference library that aided us to refine our morphology-based identification framework (Tiede et al., 2016b; Domínguez et al., 2016; Tiede et al., 2017). For analysis, ant species collected in the two seasons were pooled.

Bird species occurrence was assessed on 30 plots in the study area under favorable weather conditions at 6:00 - 10:00 and 16:00 - 18:00 (UTC-5:00) between 2000 and 2002 (Paulsch and Müller-Hohenstein, 2008; Paulsch and Wallis, 2016). Standardized point counts (30 min) were repeated 12 times on each plot and combined with mist-netting data to comprehensively assess bird assemblages.

3.2.3 Diversity measures

For each taxon, we calculated species richness and aspects of species turnover with the following considerations. In species-rich communities, the number of species recorded in a sample depends on the number of sampled individuals (Colwell et al., 2012). The number of moth individuals attracted to light traps, for example, depends on light conditions and temperature (Beck et al., 2011). In the study area, the number of tree individuals varied considerably across plots, which led to a bias in observed species numbers. Thus, we estimated species richness for trees and all moth taxa at a coverage of 70% of the expected total species richness using rarefaction and extrapolation as recommended by Colwell...
et al. (2012) using the function estimateD in the R package iNext (Hsieh et al., 2016; R Core Team, 2016, for statistics see Appendix B Table B.1).

We characterized the change in species composition (β-diversity) across plots for each taxon using ordination. For each taxon, we calculated pair-wise Bray-Curtis dissimilarities using presence/absence data between all plots for which a species list of the respective taxon was available. Subsequently, we performed NMDS using the R function metaMDS in vegan. The stress value for two-dimensional ordinations of all taxa was < 0.2 (Appendix B Table B.2), which indicates an appropriate NMDS solution. We extracted the site scores for the two dimensions. Both dimensions (NMDS I and NMDS II) are a numerical measure of the compositional change in the community across all plots.

### 3.2.4 Preprocessing of multi-spectral orthophotos

Within the framework of the Ecuadorian program ‘National System of Information on Rural Lands and Technological Infrastructure’, a digital elevation model (DEM) with a spatial resolution of 3 m and ortho-rectified images with a resolution of 0.3 m and four spectral bands (red, green, blue, near-infrared) were recorded during three flight campaigns in 2010 and 2011 under favorable weather conditions in our study area (Ministerio de Agricultura, Ganadería, Acuacultura y Pesca; Proyecto Sistema Nacional de Información y Gestión de Tierras Rurales e Infraestructura Tecnológica – SIGTIERRAS). We resampled the DEM and orthophotos to 6 m spatial resolution to match the size of canopy tree crowns. However, mountains and the airborne multi-spectral sensor, which flew at a low height, both casted shadows that appeared in orthophotos. We thus had to mask out those regions where shadows could not be corrected owing to a low contrast of the sensor. These areas were detected using the hillshade tool in the ‘spatial analyst’ of ArcGIS (version 10.3). Furthermore, we corrected the orthophotos topographically in a Java environment following Curatola Fernández et al. (2015). A problem of our procedure is the time lag between taxon sampling and image recording. The change of tree communities is, however, generally slow, and the protected forest habitats investigated (Podocarpus National Park and Reserva Biológica San Francisco) did not substantially change between the sampling of birds in 2000–2002 and the recording of orthophotos in 2010–2011 (Curatola Fernández et al., 2015; Thies et al., 2014). In addition, all in-

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>Year</th>
<th>Method</th>
<th>Total no. of species</th>
<th>Min. no. of species per site</th>
<th>Max. no. of species per site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>2007, 2008</td>
<td>Counts on 20 m × 20 m</td>
<td>443</td>
<td>12</td>
<td>41</td>
</tr>
<tr>
<td>Geometridae</td>
<td>1999, 2000</td>
<td>Light traps</td>
<td>1,223</td>
<td>135</td>
<td>296</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>1999, 2000</td>
<td>Light traps</td>
<td>753</td>
<td>62</td>
<td>315</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>1999, 2000</td>
<td>Light traps</td>
<td>443</td>
<td>34</td>
<td>149</td>
</tr>
<tr>
<td>Ants</td>
<td>2014</td>
<td>Nutrient baits</td>
<td>88</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Birds</td>
<td>2000, 2001, 2002</td>
<td>Point counts (30 min)</td>
<td>147</td>
<td>14</td>
<td>46</td>
</tr>
</tbody>
</table>
vestigated plots of the older sampling campaigns plotted on the digital orthophotos were situated in what appeared to be intact forest, and a visual interpretation of the digital orthophotos of the study area did not reveal disturbances (e.g., caused by landslides) or human-induced habitat changes, such as deforestation.

We used indicators derived from topography and textural information of RS images for characterizing topographical complexity, vegetation structure and habitat heterogeneity. For the indicators derived from topography, we used three topographical metrics, namely the DEM, the slope calculated in degrees (SLOPE), and a Topographical Position Index (TPI) using the raster package in R (Table 2; Hijmans et al., 2015). The TPI was calculated according to Wilson et al. (2007) with a surrounding of 17 pixels; this TPI compares the elevation of a pixel with the mean elevation of its environment.

For the indicators derived from textural information, we used the near-infrared (NIR) band and two vegetation indices, namely the normalized difference vegetation index (NDVI), which is based on the red and NIR bands, and an approximation of the anthocyanin reflectance index (ARI), which is based on the red and blue bands of the multi-spectral orthophotos (Table 3.3).

We chose these indices to account for different vegetation properties that might shape the habitats of our study sites. Both the NIR band and the NDVI are often used as a proxy for biomass or to account for primary productivity (Huete et al., 1997). ARI accounts for the accumulation of anthocyanin pigments in leaves (Gitelson et al., 2001; Sims and Gamon, 2002). Based on the NIR band, NDVI, and ARI, we calculated texture metrics derived from the gray-level co-occurrence matrix using the glcm package in R (Zvoleff, 2015). Among various texture statistics, we applied the statistics 'mean', 'correlation', and 'entropy' because we identified them as the most uncorrelated ones (Haralick, 1979; Appendix B Table B.3). We used a moving window size of 17 pixels × 17 pixels matching a surrounding of 102 m × 102 m to account for textural information within a relatively broad spatial scale. The image textural approach resulted in nine texture metrics. We also calculated texture
metrics derived from a second moving window (3 pixels × 3 pixels, 18 m × 18 m) to account for differences among taxa with regard to the spatial scale of their habitat demands, subsequently used as additional metrics in a second model approach (Appendix B Table B.5).

All topographical and texture metrics were extracted for the corresponding plots and polygons, respectively, of each taxon. To avoid random noise (pixel values with no relation to the image scene) in predictor images, we extracted the mean of all spatial predictor variables within polygons for ant samples, and we extracted the mean of all spatial predictor variables within a buffer of 10 m around each sample point for the remaining taxa.

3.2.6 Statistical approach

For the core analysis, we performed partial least-squares (PLS) regressions to model diversity measures for each taxon using textural and topographical metrics. PLS regressions were developed for situations where a low number of samples has to be modeled against a large number of inter-correlated predictor variables (Carroll et al., 2009). To reduce the number of predictors and to deal with multi-collinearity, PLS regression models derive latent vectors from the predictors that explain the maximum variance of the response variable. Even though this regression is a reliable method, the generation of latent vectors by the PLS algorithm is unfortunately a black-box procedure. Since the structure of latent vectors would change if predictors are added or removed, no precise information on the explained variance of single predictors can be made. Therefore, the possible variety of statistically reasonable predictor combinations could lead to different conclusions about the relationship of species diversity and habitat indicators. We chose those models with the lowest prediction error and subsequently discussed the ecological meaning of these models using the most important predictors.

We used the autopls package in R (Schmidtlein et al., 2012), which implements a variable selection to reduce the size of the set of predictor variables. At the same time, the variable selection reduces computation time, and improves both model performance and interpretability (Andersen and Bro, 2010). This optimization procedure was based on a filter combining significance of predictors estimated by jackknifing and variable importance in the projection (VIP). The VIP scores are based on the weighted sums of the absolute regression coefficients across the number of latent vectors (Chong and Jun, 2005).

Following our proposed aims, we fitted models for species richness and the two NMDS axes characterizing species turnover to compare their

<table>
<thead>
<tr>
<th>Topographical metrics</th>
<th>Abbr.</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>DEM</td>
<td>Digital elevation model</td>
</tr>
<tr>
<td>Slope</td>
<td>SLOPE</td>
<td>Gradient of DEM in degree</td>
</tr>
<tr>
<td>Topographical Position Index</td>
<td>TPI</td>
<td>After Wilson et al. (2007) with a surrounding of 17 pixels</td>
</tr>
</tbody>
</table>
Table 3.3: Optical band and vegetation indices used as base layers for the calculation of texture metrics (see Appendix B Table B.3).

<table>
<thead>
<tr>
<th>Optical metrics</th>
<th>Abbr.</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near infra-red band</td>
<td>NIR</td>
<td>Pre-processed near infra-red band</td>
</tr>
<tr>
<td>Normalized difference vegetation</td>
<td>NDVI</td>
<td>(NIR - red) / (NIR + red), sensitive to chlorophyll pigments</td>
</tr>
<tr>
<td>index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthocyanin reflectance adjusted</td>
<td>ARI</td>
<td>(red - blue) / (red + blue), sensitive to anthocyanin</td>
</tr>
</tbody>
</table>

performance using the leave-one-out (LOO) validated $R^2$. For all models, we tested the regression residuals for spatial autocorrelation using Moran’s I with a neighborhood of five neighbors, but we did not find a significant autocorrelation of the residuals in any of the models ($p < 0.05$). To identify both the most important predictors for each model and an overall trend among species richness and turnover, we used VIP values as a measure of predictor importance, and regression coefficients as a measure of predictor influence (Chong and Jun, 2005). In general, special emphasis is put on the VIP values greater than one, since the average of squared VIP scores equals one (Mehmood et al., 2012). However, to consider a model-specific VIP cutoff and to facilitate the interpretation of predictor variables, we placed emphasis on all VIP scores greater than the third quartile of all VIP observations in each model.

3.3 Results

3.3.1 Species diversity measures

In general, the composition of species along NMDS axes is often related to species richness. However, we found significant correlations only between species richness and NMDS I for pyraloid moths ($r = 0.82$, $p < 0.01$) and a correlation between species richness and NMDS II for ants ($r = 0.66$, $p < 0.01$). Since the three moth taxa were sampled on the same plots, we were able to test for correlations between diversity measures of these taxa. Measures of species richness showed no significant correlations between moth taxa, but correlations were found for NMDS I ($r>|0.95|$ in all three cases; $p < 0.01$) and for NMDS II ($r>0.81$; $p < 0.01$).

3.3.2 Predictive power of diversity models

The predictability of species richness highly varied among taxa, ranging from 4% of explained variance for arctiinae moths to 78% for ants (Table 3.4). When image textures from a different window size ($3 \times 3$ pixels; $18 m \times 18 m$) were used as additive predictors (Appendix B Table B.5), the predictive power of species richness of trees and birds increased significantly from 12% to 22% and from 13% to 27%, respectively. In contrast to species richness, the models of NMDS I showed universally high LOO-validated $R^2$ values for all six taxa, ranging from 0.64 (birds) to 0.98 (trees; Table 3.4). NMDS II showed more variation in model performance among taxa than NMDS I, and ranged from 33% of explained variance for trees to 74% for pyraloid moths (Table 3.4).
3.3 Results

Table 3.4: Results of partial-least squares regressions with implemented backward selection and leave-one-out (LOO) cross validation for the three diversity measures and all studied taxa using a set of spatial predictors (n = 12). LV = latent vector.

<table>
<thead>
<tr>
<th>Species diversity measure</th>
<th>Taxonomic group</th>
<th>$R^2$</th>
<th>LOO $R^2$</th>
<th>No. of LVs</th>
<th>No. of predictors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species richness</td>
<td>Trees</td>
<td>0.30</td>
<td>0.12</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>0.84</td>
<td>0.72</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Geometridae</td>
<td>0.75</td>
<td>0.57</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Arctiinae</td>
<td>0.40</td>
<td>0.04</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>0.85</td>
<td>0.78</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>0.33</td>
<td>0.13</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>NMDS</td>
<td>Trees</td>
<td>0.99</td>
<td>0.98</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>0.96</td>
<td>0.94</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Geometridae</td>
<td>0.96</td>
<td>0.90</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Arctiinae</td>
<td>0.97</td>
<td>0.89</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>0.94</td>
<td>0.89</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>0.78</td>
<td>0.64</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>NMDS II</td>
<td>Trees</td>
<td>0.41</td>
<td>0.33</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>0.95</td>
<td>0.74</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Geometridae</td>
<td>0.94</td>
<td>0.71</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Arctiinae</td>
<td>0.78</td>
<td>0.64</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>0.74</td>
<td>0.63</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>0.59</td>
<td>0.45</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

3.3.3 Predictor importance and influence

The species richness of both pyraloid moths and ants increased with decreasing DEM values and, in addition, that of ants increased with increasing 'entropy' texture of ARI, while that of pyraloid moths increases with decreasing 'mean' texture of ARI (Figure 3.2). By contrast, DEM was not included in the models of geometrid moths and arctiinae moths, and it showed low importance in the models of trees and birds (VIP values < 3rd quartile, Figure 3.2). The species richness of these taxa increased with increasing 'mean' and 'entropy' texture metrics. Tree species richness was negatively associated with the TPI and positively associated with the 'entropy' statistics of the NIR band (Figure 3.2).

Since NMDS I and II are based on ordinations, the direction of predictor influence is arbitrary. NMDS I of all taxa except ants was predicted by DEM; for pyraloid moths, DEM and TPI were the only predictors (Figure 3.2). NMDS I of bird species was also predicted by SLOPE and the texture 'correlation' of ARI, whereas NMDS I of trees, geometrid moths, and arctiinae moths was also predicted by additional 'mean' and 'entropy' texture statistics. For trees, NMDS II was predicted by TPI, while DEM and 'mean' and 'entropy' texture statistics together predicted NMDS II of ants (Figure 3.2). NMDS II of all
ectotherms animals was predicted by 'entropy' texture metrics of NDVI.

3.4 Discussion

Our models revealed considerable variation in the ability of habitat indicators to predict species richness across the six taxa studied. Four out of six taxa were predicted by RS texture metrics, largely independent of elevation. Our models also revealed the high predictability of species turnover assessed by NMDS ordination of all six taxa. A combination of elevation and taxon-specific texture metrics explained species turnover along an elevational gradient, and texture metrics, particularly 'entropy' texture metrics, explained species turnover along more subtle, local changes in habitat structure.

3.4.1 Species richness and elevation

Considering the well-documented response of ectotherms to temperature (e.g., McCain and Grytnes, 2010), it is not surprising that the number of species of pyraloid moths and ants decreased with increasing elevation. Consequently, elevation is a good proxy for the ther-
3.4 Discussion

In contrast to the decrease of species richness of pyraloid moths, models for ants slightly benefited from the inclusion of additional texture metrics. In combination with elevation, textural information provided further facets of the environment to model additional variance in species richness, such as 'entropy' statistics of ARI for ants. ARI accounts for the accumulation of anthocyanin pigments in leaves (Gitelson et al., 2001; Sims and Gamon, 2002). Anthocyanin pigments are generally higher in senescent leaves but occur also as protection against herbivory mostly in young leaves (Karageorgou and Manteas, 2006), and might protect against UV radiation (Caldwell, 1981; Steyn et al., 2002). Generally, tropical mountain regions have high UV-B levels, which increase with altitude and canopy openness (Flenley, 2011). 'Entropy' statistics account for the degree of disorder (Haralick, 1979) and thus high values of 'entropy' ARI depict a higher vegetation heterogeneity somehow related to anthocyanin pigments in canopy leaves. In addition, 'entropy' ARI decreased with a higher C/N ratio in plant tissues (r = -0.55, p < 0.01; Appendix B Table B.4), which is a proxy for N availability in forest soils that is related to nutrient use efficiency of plants (Wolf et al., 2011). Therefore, this texture metric might be an important indicator of nutritional quality of foliage for herbivores.

3.4.2 Species richness and habitat structure

Species richness of trees, geometrid and arctiinae moths, and birds showed no significant response to elevation, and models using 'topography only' as predictor variables (Appendix B Table B.6) were weak or even failed to explain the variability of species richness (Appendix B Table B.6). This indicates that information from topography cannot predict the diversity of these taxa successively, which confirms the findings regarding species richness of geometrid moths in Brehm et al. (2016).

Species richness of trees, geometrid moths, arctiinae moths and birds was better predicted by textural predictors, such as 'mean' NIR for geometrid moths and birds, and 'entropy' NDVI/ARI for arctiinae moths. For our study area, Pearson correlations suggested that 'mean' and 'entropy' texture metrics were equal or even more highly related to site-based forest productivity data (r = 0.40—0.78, p < 0.01) than topographical metrics (r = 0.38—0.64, p < 0.01; Appendix B Table B.4). We suppose that 'entropy' statistics of the NIR band and NDVI thus depict a high diversity of vegetation types. This habitat heterogeneity, which is probably related to a high structural complexity in ravine forests or forest edges, might support more coexisting species.

Particularly the predictability of geometrid moths that explains 57% of species richness should be noted because earlier studies that assessed only the effects of temperature or elevation revealed much less predictive power (max(|r|) = 0.29, p > 0.24, Brehm et al. 2016; R² = 0.0243, p>0.488, Fiedler et al. 2008). As suggested
in Stein et al. (2014), habitat heterogeneity might be more important for certain herbivores than climatic or topographical heterogeneity since different clades of moth species have developed distinct biological features to cope with environmental factors such as temperature (Braga and Diniz, 2015). By contrast, arctiinae moths comprise a speciose clade of which a large amount of caterpillar species are not classical herbivores as they feed on dead or withered leaves or epiphylls (Bodner et al., 2015; Seifert et al., 2016). Therefore, it is not surprising that neither topographical nor texture metrics as proxies for habitat heterogeneity (dominated by tree crowns) were able to model the species richness of arctiinae moths successively.

Although the predictability of species richness of trees and birds was low, we showed that it significantly increased when texture metrics of multiple window sizes were used (Appendix B Table B.5), as recommended by Mairota et al. (2015) to match the approximate scale of activity (e.g., home ranges) among species of the targeted taxonomic group. Birds were sampled only along a short stretch of the elevational gradient (700 m; Figure 3.1), which might cause a higher dependence of species richness on habitat heterogeneity than on topography. However, general patterns of bird species richness remain unclear because of contrasting results obtained in tropical forests along elevational gradients, ranging from increasing bird species richness with higher habitat heterogeneity (Wallis, 2016), to decreasing species richness with increasing elevation (Jankowski et al., 2013), to a hump-shaped relationship between elevation and species richness (Herzog et al., 2005). Our models of tree species richness were associated with negative TPI values that represent locations that are on average lower in elevation than its surrounding and that have lower values of ‘correlation’ NDVI, which indicated areas of uniform NDVI values. Thus, in accordance with Werner and Homeier (2015), higher tree species richness might occur in valleys where the canopy is higher and denser (= high leaf area index) representing higher soil nutrient availability compared to upper slopes and ridges. The low to moderate predictive power of tree species richness found in this study and in earlier studies that assessed multispectral RS metrics (e.g., Fricker et al., 2015) lead to the need for information on more subtle variance in either habitat characteristics or the spectral response of woody plant species in the canopy. Variables derived from RS information with higher spectral resolution, as recommended by Asner and Martin (2011), have been successfully used to model tropical tree alpha-diversity (Féret and Asner, 2014; Schäfer et al., 2016). For example, Vaglio Laurin et al. (2014) have shown that hyperspectral data explain up to 84.9% of the alpha-diversity of upper canopy trees in a West African forest, whereas common vegetation indices yield poor results. We therefore propose that pigment-related hyperspectral data would be more successful in modeling tree species richness than habitat heterogeneity assessed with image texture metrics.

### 3.4.3 Species turnover along environmental gradients

Models for species turnover showed far higher predictive power among all taxa than models
for species richness. The scores along the first NMDS axis were strongly associated with elevation. By contrast, the scores of the second NMDS axis showed a response to the elevational gradient only for ants. Species have evolved specific adaptations to climatic zones that influence their distribution (Angilletta Jr., 2009, for ants, e.g. Bishop et al. 2015; for geometrid moths, e.g., Brehm et al. 2003; for other moths species, e.g., Fiedler et al. 2008). Thus, it is not surprising that the elevational gradient is crucially important for the compositional change (Guerin et al., 2013). It is obvious that other factors besides elevation influenced the composition and turnover of the investigated assemblages, but our results underline the importance of elevation for shaping an important part of species turnover of almost all taxa, from primary producers to herbivores, detritivores, and predators. The scores of the second dimension of NMDS ordination (for ants, the first NMDS axis) indicated a change in species composition predicted by texture metrics that characterize local habitat heterogeneity. Assemblages of all taxonomic groups of moths and ants changed with primary productivity and habitat heterogeneity (‘entropy’ NDVI), or in the case of arctiinae moths, with variation in ARI, which is related to the C/N ratio in understory plant tissues. This might explain the small-scaled compositional change in species assemblages along forest edges or glades to forests with a higher canopy cover, as observed in the study area for a range of organisms (Brehm et al., 2003; Hilt and Fiedler, 2006; Hilt et al., 2006; Wallis, 2016). Thus, the second ordination axis appears to account for compositional variation of communities along a habitat structural gradient, largely independent of elevation.

3.4.4 Conclusions

Indicator applications differed among both the taxa and diversity measures studied owing to distinct variation in patterns of species distribution. Species richness models of trees, geometrid moths, and birds highly benefited from integrating RS-based texture metrics. By contrast, topographical metrics sufficiently modeled species richness of pyraloid moths and ants. Models of species turnover can be constructed across taxonomic groups with far higher predictability than models of species richness. The two ordination axes effectively separated compositional change along the elevational gradient (NMDS I) from more subtle, changes in habitat structure, which were surrogated by RS-based texture metrics (NMDS II). However, we have to stress that the interpretation of causal relationships between RS texture metrics and species diversity is sometimes difficult and that field based inventories are still necessary to calibrate RS proxies. In addition, airborne RS missions are often not repeated and thus lack a temporal aspect, which is essential for recurrent monitoring. We therefore recommend to obtain habitat indicators from upcoming missions of satellites with a fine to moderate spatial and temporal resolution (e.g., Sentinel-2 mission; Drusch et al., 2012), which would allow the generation of cloud-free composites for different time spans. Hence, textural information derived from future satellite missions as habitat indicators would allow monitoring species diversity of a range of taxa successively in space and time.
Acknowledgements

This study was conducted in the framework of the German Research Foundation (DFG) projects FI 547/5-1 to 5-3 and the Research Units FOR 402, FOR 816 (subproject A1 HO3296/2-2), and PAK 823-825 (subproject C2 BE1780/34-1, BR1293/11, FA925/7-1, ZI698/8-1). We are grateful to the DFG (PAK 825/1) for funding our project, to the Ecuadorian Ministry of the Environment (MAE) for permission to conduct research, and to the foundation Nature and Culture International (NCI) for logistic support. We acknowledge the support of all our research assistants who helped with field work and sample processing. The Ministerio de Agricultura, Ganadería, Acuacultura y Pesca kindly provided the SIGTIERRAS orthophotos, which were the basis of our approach. We thank Karen A. Brune for linguistic revision of the manuscript. Special thanks go to two anonymous reviewers for constructive comments.

References

References


Chapter 3 Modeling species richness and turnover of six taxa by topographical and textural metrics


Féret, J.B., Asner, G.P., 2014. Mapping tropi-
References


Hsieh, T.C., Ma, K.H., Chao, A., 2016. iNEXT: Interpolation and extrapolation for species diversity.


References


Chapter 3 Modeling species richness and turnover of six taxa by topographical and textural metrics


Vaglio Laurin, G., Chan, J.C.W., Chen, Q., Lind-


Zvoleff, A., 2015. Glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R.
Chapter 4

Modeling tropical montane forest biomass, productivity and canopy traits with multispectral remote sensing data

Christine I. B. Wallis, Jürgen Homeier, Roland Brandl, Nina Farwig, Jörg Bendix

Submitted to Remote Sensing of Environment
Chapter 4  Modeling forest biomass, productivity and canopy traits with multispectral data

Abstract

Remote sensing allows modeling of variables related to spatial patterns of carbon stocks and fluxes. However, at the landscape scale, most studies conducted so far are based on airborne remote sensing data such as Airborne Laser Scanning (ALS) surveys. In contrast, multispectral satellite data with an appropriate life span and repetition rate are limited to modeling biomass, productivity or canopy traits at the landscape scale due to their spectral or spatial resolutions. However, image textures as proxies for the habitat structure might improve current approaches using multispectral satellite data. Here, we investigated the relationship between multispectral remote sensing metrics at medium spatial resolution and forest biomass, its productivity and selected canopy traits of a tropical montane forest. To train and validate the models, in situ data were sampled in 54 permanent plots in forests of southern Ecuador distributed within three study sites at 1000 m, 2000 m and 3000 m a.s.l. Topographical, spectral and textural metrics were derived from the Aster Elevation Model and a Landsat-8 OLI image. We used partial least squares regressions to model and predict all response variables. Along the whole elevation gradient, biomass and productivity models explained 31%, 43%, 69% and 63% of the variance in aboveground biomass, annual wood production, fine litter production and aboveground net primary production, respectively. Regression models of canopy traits measured as community-weighted means explained 62%, 78%, 65% and 65% of the variance in leaf toughness, specific leaf area, foliar N concentration, and foliar P concentration, respectively. Models at single study sites hardly explained variation in aboveground biomass and the annual wood production indicating that these measures are mainly determined by elevation. In contrast, the models of fine litter production and canopy traits explained between 8% and 85% of the variance, depending on the study site. As spectral metrics, in particular, a vegetation index using the red and green bands provided complementary information to topographical information, and the model performance in terms of leaf toughness, biochemical canopy traits and related fine litter production improved. Although textural metrics were correlated with forest productivity and canopy traits, their benefit in addition to topographical metrics was marginal. Our findings therefore revealed that fine litter production and canopy traits are driven by local changes in vegetation edaphically induced by topography. We conclude that multispectral remote sensing based on operational data on the landscape scale including spectral metrics (in particular those not related to topography) is an important tool to model underlying drivers of carbon-related variables, such as fine litter production and biochemical canopy traits, in topographically and ecologically complex tropical montane forest.

4.1 Introduction

Forest primary production is a key for understanding the interactions between climate change and the carbon cycle (Houghton et al., 2009). Spatially explicit estimates of the carbon stocks and fluxes in forests are a prerequisite for land use management to mitigate rising carbon emissions. Carbon cycling in forests might be supported by a quantitative understanding of biomass, its productivity and key traits of the species involved (Reich et al., 2012). Tropical rainforests are the most productive terrestrial ecosystems and account for the highest terrestrial net and gross primary production per unit area (Beer et al., 2010; Moore et al., 2018). In particular, tropical montane forests store higher amounts of biomass along their
steep slopes than expected (Spracklen and Riglato, 2014). The measurements of the pools and fluxes of the carbon cycle in these topographically and ecologically complex ecosystems are very labor-intensive. Thus carbon cycling in tropical montane forests is poorly understood compared to tropical lowland and extratropical forests (Anderson-Teixeira et al., 2016).

Remote sensing has been frequently used as a tool to map and monitor variables related to carbon sequestration in space and time (Timothy et al., 2016). Most of these studies are based on airborne remote sensing sources, such as costly Airborne Laser Scanning (ALS) or surveys using hyperspectral imaging (Fassnacht et al., 2014; Sinha et al., 2015). In particular, the use of multispectral satellite data in tropical rainforests is rare since frequent cloud coverage hampers the collection of cloud-free images. Notwithstanding, operational satellite data are advantageous relative to airborne data since their life span and repetition rate enables long-term monitoring. Moreover, medium-resolution satellite images such as data from the Landsat or Sentinel missions are freely available (Popkin, 2018).

Estimates of tropical aboveground biomass (AGB) have been derived from medium resolution satellite images with differing predictive power (Lu et al., 2014; Timothy et al., 2016). Landsat derived metrics explained 25% - 30% of AGB in tropical forests of Brazil, Malaysia and Thailand (Foody et al., 2003) and between 43% and 72% of the AGB of commercial forests in South Africa (Dube and Mutanga, 2015a). At continental to global scales, multispectral data has been intensively used to model the productivity of forests (Song et al., 2013). In particular, global gross and net primary production (NPP) products from the Moderate-resolution Imaging Spectroradiometer (MODIS) are widely used to monitor primary production at broad scales (Robinson et al., 2018; Turner et al., 2006; Yin et al., 2017; Zhang et al., 2009). An improved product of gross primary productivity and NPP at finer resolutions for the United States has been developed by Robinson et al. (2018) using Landsat images. These estimations, however, have limited predictive power in tropical forest regions at the landscape scale. Therefore, an improved understanding of the relationship between multispectral remote sensing data in addition to tropical forest biomass and productivity is necessary at the landscape scale.

As proposed by Finegan et al. (2015) and Coops (2015), estimates of forest productivity can be also derived through remotely sensed foliar leaf traits. Recent studies investigated the composition of stands with respect to functional leaf traits such as specific leaf area (SLA) or foliar leaf nutrients as potential driver of tropical AGB and NPP (e.g., Mercado et al., 2011; Reich et al., 2012). Fyllas et al. (2017) suggested that spatial variations in canopy traits can surrogate the spatial variation in productivity along an elevation gradient in the Amazon-Andes. In particular, the SLA, leaf nitrogen concentration and the force to tear the leaf are important predictors for annual biomass increments in three different tropical rainforests (Finegan et al., 2015). Accordingly, a variety of remote sensing studies have focused on recording leaf traits to access ecosystem processes and services at various scales (Martínez et al., 2016). To match the scale of field based leaf traits and remote sensing met-
rics, leaf traits are averaged and weighted per forest stand (community-weighted mean, CWM; Homolová et al., 2013).

Similar to AGB and NPP, CWMs of leaf traits (hereafter canopy traits) are generally retrieved from hyperspectral sensors (Homolová et al., 2013). A few studies have demonstrated that multispectral satellite data at medium spatial resolutions can be used to model morphological canopy traits such as LAI or SLA (e.g., Ali et al., 2017; Middinti et al., 2017) in addition to biochemical canopy traits (Lepine et al., 2016; Shiklomanov et al., 2016). Canopy traits, however, are often intercorrelated (Reich et al., 1999; Wright et al., 2005, 2004). Maximizing photosynthesis codetermines also the properties of the leaves that influence the spectral reflectance, such as foliar chlorophyll, N and P concentrations (Baraloto et al., 2010; Falster and Westoby, 2005). This led us to the hypothesis that photosynthesis-related spectral metrics such as the enhanced vegetation index (EVI) can predict canopy traits even if they are missing a specific spectral signal or generally cannot be retrieved using coarse-spectral-resolution data.

Regarding forest stands with a high biomass, prior studies often found a saturation effect of the spectral signal with high amounts of biomass or canopy traits (Asner and Martin, 2008; Huete et al., 1997; Lu et al., 2012; Zhao et al., 2016). It has been shown that the inclusion of image textures used in addition to the spectral signal improves the estimation of variables characterizing forest structure (Wood et al., 2012; Wulder et al., 1998) and aboveground biomass (Kelsey and Neff, 2014; Maack et al., 2015). In general, image textures represent the structure of an image considering the spatial organization of spectral information. At medium spatial resolutions, image textures have been used not only to distinguish patterns of among-habitat vegetation structure (Wood et al., 2012) but also to predict forest biomass (Safari and Sohrabi, 2016). According to their success within vegetation studies, textural information extracted from remote sensing images might therefore address the relationships of canopy structure with forest biomass, productivity, and canopy traits more effectively than point-based vegetation indices.

Here, we investigated the potential of multispectral remote sensing data to predict forest AGB, productivity and selected canopy traits in tropical montane forests of southern Ecuador. As we assume differences in the predictability of canopy production and the production of woody biomass, we measured the annual wood production ($A_{GB}$), fine litter production (FLP) and the sum of both, aboveground NPP ($NPP_a$), in permanent forest plots. In the same plots we sampled canopy leaves to calculate CWMs of leaf toughness (the force required to punch a leaf), the SLA, and foliar N and P concentrations as canopy traits. Remote sensing metrics were derived from Aster Elevation Model data and a Landsat-8 OLI image, including topographical information, vegetation indices and their textural information. We used partial least squares regression (PLSR) to model AGB, forest productivity, and canopy traits along an elevation gradient and at individual elevation sites.
4.2 Data and Methods

4.2.1 Study area

The study was conducted in an area of south-eastern Ecuador known for its high species richness (Brehm et al., 2008). We focused on three study sites: Bombuscaro (lat: -78.97204, long: -4.119909), San Francisco (lat: -79.07320, long: -3.975525) and Cajanuma (lat: -79.17794, long: -4.111812), which follow an elevational gradient (1000 m - 3000 m a.s.l.) and thus gradients of abiotic variables (Figure 4.1 a, b). The three study sites partly overlap with the Reserva Biológica San Francisco and Podocarpus National Park. The climate is humid throughout the year, with a peak in precipitation from June to August and a drier period from December to February. The mean annual precipitation ranges between 2200 mm at an elevation of 1000 m a.s.l. (Bombuscaro) and 4500 mm at 3000 m a.s.l. (Cajanuma; Beck and Kottke, 2008; Bendix et al., 2006). The mean annual air temperature decreases from $20^\circ$C at $\sim1000$ m a.s.l. to $15.5^\circ$C at $\sim2000$ m a.s.l. to $9.5^\circ$C at $\sim3000$ m a.s.l. (Beck and Kottke, 2008; Bendix et al., 2008).

Due to the high vertical and horizontal structural diversity of these forests, various forest types have been recognized in the study area: ev-
ergreen premontane forest (≤1300 m a.s.l., Bombuscaro site), evergreen lower montane forest (between 1300 m and 2100 m a.s.l., San Francisco site), and evergreen upper montane forest (>2100 m a.s.l., San Francisco and Cajanuma sites), with characteristic subtypes along valleys, ravines, ridges, and anthropogenic replacement systems (Homeier et al., 2008). These forest types differ in terms of species richness, floristic composition, and structural characteristics, which in turn should affect their spectral and textural responses. The area is characterized by relatively nutrient-poor soils on metamorphic schists and sandstones. Nutrient supplies of nitrogen and phosphorus are slightly better at lower elevations than on upper slopes and at upper elevations (Werner and Homeier, 2015; Wolf et al., 2011). The decreasing nutrient availability is reflected by decreasing biomass and productivity with increasing elevation (Leuschner et al., 2013; Wolf et al., 2011) and with the topographical gradient from lower to upper slope position (Werner and Homeier, 2015).

4.2.2 Data regarding forest productivity and functional leaf traits

We sampled data regarding aboveground woody biomass, forest productivity and functional leaf traits in 54 permanent plots, with 18 plots in each of the three study sites (Figure 4.1 b, c). Each plot was 20 m × 20 m. GPS records were taken in the center of each plot. All plots were located in mature old-growth forest without visible signs of human disturbance within the plots or in their direct surroundings. We calculated AGB and AGB, from two tree inventories of all stems ≥10 cm diameter at breast height (dbh) in 2008 and 2009 (Leuschner.2013). To calculate AGB, we used the allometric equation for tropical wet forests proposed by Chave et al. (2005), which is based on stem diameter, wood specific gravity (WSG) and tree height as parameters. WSG estimates for all trees were obtained with a Pilodyn 6J wood tester (PROCEQ SA, Zuerich, Switzerland) following the method of Chave et al. (2008).

Total litterfall was collected monthly (biweekly at 1000 m) during one year (5/2008 – 5/2009) using six 60 cm × 60 cm litter traps within each plot. The samples were dried at 60°C and weighed. We calculated NPPa as the sum of annual AGBi and annual FLP (Appendix C Table C.1). From the same permanent plots, we collected canopy leaves from a minimum of ten randomly selected trees (≥ 10 cm dbh; $\bar{x}$=14.2±2.4 per plot) and determined the SLA, leaf toughness, foliar N and foliar P concentrations (for detailed methods see Báez and Homeier, 2018). The canopy traits (CWMs of leaf traits) represent the trait values of an average tree within the community. We calculated the CWMs of each trait for each plot by weighting tree species contributions to the plot mean by their basal area (Appendix C Table C.1; Lavorel et al., 2008).

4.2.3 Remote sensing predictors

A Landsat-8 scene recorded on 11/20/2016 and four Aster elevation images were used to calculate predictor variables for modeling. The Landsat scene was acquired as level 2 surface reflectance data preprocessed by USGS, including geometric, radiometric, and atmospheric calibrations. We further performed a topographic correction in a Java environment following Cu-
4.2 Data and Methods

Table 4.1: Topographical, spectral and textural metrics. Texture statistics derived from gray level co-occurrence matrices were calculated based on the three spectral metrics. In the moving window approach, which shifted in all directions, a 3 pixel × 3 pixel window was employed. In total, three topographical, three spectral and 6 textural metrics were calculated. We removed metrics that returned NAs. In total 11 metrics remained. NIR = near-infrared.

<table>
<thead>
<tr>
<th>Predictor set</th>
<th>Metrics</th>
<th>Abbr.</th>
<th>Metric description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topographical</td>
<td>Elevation</td>
<td>DEM</td>
<td>Digital elevation model</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>Slope</td>
<td>Gradient of DTM in degree</td>
</tr>
<tr>
<td></td>
<td>Topographical Position Index</td>
<td>TPI</td>
<td>After Wilson et al. (2007), with a surrounding of 11 pixels</td>
</tr>
<tr>
<td>Spectral</td>
<td>Enhanced vegetation index</td>
<td>EVI</td>
<td>( \frac{2.5 \times (\text{NIR band} - \text{red band})}{(\text{NIR band} + 6 \times \text{red band} - 7.5 \times \text{blue band} + 1)} ) *</td>
</tr>
<tr>
<td></td>
<td>Red-Green vegetation index</td>
<td>RGVI</td>
<td>( \frac{\text{red band} - \text{green band}}{\text{red band} + \text{green band}} )</td>
</tr>
<tr>
<td></td>
<td>Red-Blue vegetation index</td>
<td>RBVI</td>
<td>( \frac{\text{red band} - \text{blue band}}{\text{red band} + \text{blue band}} )</td>
</tr>
<tr>
<td>Textural</td>
<td>Entropy</td>
<td>EN</td>
<td>( EN = \sum_{i,j=0}^{N-1} p_{i,j}(-\ln p_{i,j}) ) **</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
<td>CC</td>
<td>( CC = \sum_{i,j=0}^{N-1} p_{i,j} \left[ \frac{(i-ME)(j-ME)}{\sqrt{VA \cdot VA}} \right] ) **</td>
</tr>
</tbody>
</table>

* with coefficients derived from the Landsat-8 product guide (Vermote et al., 2016)

** with \( p_{i,j} = V_{i,j} \sum_{i,j=0}^{N-1} V_{i,j} \), where \( V_{i,j} \) is the value in cell \( i, j \) and \( N \) is the number of rows or columns; with \( ME = \sum_{i,j=0}^{N-1} (p_{i,j}) \) and \( VA = \sum_{i,j=0}^{N-1} p_{i,j}(i-ME)^2 \).

ratola Fernández et al. (2015). All spatial data were resampled to 30 m per pixel and projected to UTM 17 South.

To characterize the topography, we used three topographical metrics, namely, elevation in m a.s.l. (elevation), slope calculated in degrees (slope), and a Topographical Position Index (TPI). The TPI was calculated following Wilson et al. (2007); this particular TPI compares the elevation of a central pixel with the mean elevation of its surrounding, defined as 11 pixels × 11 pixels. As spectral proxies for greenness and tree health conditions, we calculated the following vegetation indices (Table 4.1): the Enhanced Vegetation Index (EVI), using the Red-Green vegetation index (RGVI; Wallis, 2016), and the Red-Blue Vegetation index (RBVI; Wallis et al., 2017; Wallis, 2016). The latter two are normalized difference indeces which have successfully been used in vegetation and biodiversity studies (Motohka et al., 2010; Wallis, 2016; Wallis et al., 2017). To account for canopy structure, we calculated the gray level co-occurrence matrix (glcm) based image textures. The glcm refers to the spatial variation in reflectance in the image. According to prior studies, image textures should be able to model tree canopy cover and foliage height (Karlson et al., 2015; Wood et al., 2012). We used
Chapter 4 Modeling forest biomass, productivity and canopy traits with multispectral data

'entropy' and 'correlation' statistics based on the three vegetation indices (Table 4.1). These two measures are known to be the most uncorrelated second-order texture measures (Haralick, 1979). Whereas 'entropy' statistics derive textures as a measure of spatial disorder with high values in homogeneous areas, 'correlation' statistics measure the linear dependency of reflectance within the glcm. We applied a 3 pixel × 3 pixel moving window size to make sure that the windows of close sampling plots did not overlap for independence reasons. The texture statistic 'correlation' sometimes returns no-data values since the software cannot treat 0 (no correlation) as the denominator in the 'correlation' equation. This was the case for 'correlation' statistics of the RGVI, which was subsequently excluded from the further analysis. Image processing of the topographical, spectral and image textural calculations was performed in R (R Core Team, 2016) using the raster package (Hijmans et al., 2015) and the glcm package (Zvoleff, 2015).

4.2.4 Statistical methods

We used partial least squares regressions (PLSR) to model and predict forest biomass, productivity and canopy traits using the proposed topographical, spectral and textural metrics (Table 4.1). We used the autopls package in R (Schmidtlein et al., 2012). We applied an optimization procedure using a filtering based on thresholds for backward selection, i.e., significance, variable importance in the projection (VIP), or both (Chong and Jun, 2005; Mehmood et al., 2012). The magnitudes of regression coefficients and predictors significance were used to identify the predictors with the highest importance. In total, we built eight regression models for the whole elevation gradient with the following response variables: AGB, AGBi, FLP, NPPa, and leaf toughness, SLA, foliar N and foliar P. We subsequently modeled all response variables at the individual study sites of Bombuscaro, San Francisco and Cajanuma to identify whether topographical, spectral and textural metrics are able to model the response variables at individual elevation levels. We compared the predictability between response variables within the whole elevation gradient and within single study sites using the leave-one-out (LOO) validated R² value (Appendix C Table C.2) and the root mean square error (RMSE). We mapped all models and compared them visually. We masked out non-forest areas using the land cover classification published in Göttlicher et al. (2009) since our models were only targeting forests.

4.3 Results

4.3.1 Models of forest biomass, productivity and canopy traits

The best subset PLSR models along the whole elevation gradient exhibited moderate to high predictive power (Appendix C Table C.2; Table 4.2): 31% explained variance in AGB, between 43% and 69% in forest productivity and between 62% and 78% of variance in canopy traits. The highest predictive power among productivity models was found for the annual FLP (LOO validated R² = 0.69) with an LOO validated root mean square error of less than 1 Mg ha⁻¹ yr⁻¹. The models of NPPa explained 63% of the variance, and AGBi explained 43% of the variance. For canopy traits, the respective models explained
4.3 Results

Figure 4.2: Observed versus predicted values for all PLSR models. Hollow points = calculated values, filled points = predicted values. The dashed line represents the 1:1 line (y=x). Aboveground biomass (AGB) in Mg ha\(^{-1}\), and variables of forest productivity in Mg ha\(^{-1}\) yr\(^{-1}\). Community-weighted means of leaf toughness in kN m\(^{-1}\), of specific leaf area (SLA) in cm\(^2\) g\(^{-1}\) and of foliar N and P in mg g\(^{-1}\). Plots for models of individual study sites are appended (Appendix C Figure C.1 and Figure C.2). AGB\(_i\) = annual wood production, FLP = annual fine litter production, NPP\(_a\) = aboveground net primary production.

62%, 78%, 65% and 65% of the variance in the CWMs of leaf toughness, SLA, foliar N and foliar P concentrations.

We found that the plots of Cajanuma were clustered with lower values in biomass, productivity and canopy traits (except for leaf toughness) than in the other two study sites, as shown by regression plots of the observed versus predicted values color-coded by study site (Figure 4.2). For the leaf toughness, the sampled plots exhibited higher observed values than for Bombuscaro and San Francisco.

The PLS regressions of individual study sites mostly failed to model forest biomass and productivity or performed poorly. Only the models of FLP were able to predict a reasonable part of observed variance within Bombuscaro at approximately 1000 m a.s.l. (LOO validated R\(^2\) = 0.61) and within San Francisco at approximately 2000 m a.s.l. (LOO validated R\(^2\) = 0.48; Table 4.2). The models of canopy traits for individual study sites were able to model variations of CWMs within all study sites. Whereas the canopy trait models for Cajanuma explained be-
Table 4.2: Results of partial-least squares regressions (PLSR) with leave-one-out (LOO) cross validation for forest biomass, productivity and community-weighted mean (CWM) trait values using a set of spatial predictors (Table 4.1). LOO RMSE values of AGB in Mg ha\(^{-1}\), forest productivity in Mg ha\(^{-1}\) yr\(^{-1}\), leaf toughness in kN m\(^{-1}\), specific leaf area (SLA) in cm\(^2\) g\(^{-1}\), foliar N and P in mg g\(^{-1}\). LOO validated RMSEN in % is the RMSE normalized by the mean value of the observed data. LV = latent vector.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Sites</th>
<th>LOO (R^2)</th>
<th>LOO RMSE</th>
<th>LOO RMSEN in%</th>
<th>No. of LV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forest biomass</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGB</td>
<td>All</td>
<td>0.31</td>
<td>64.95</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.43</td>
<td>1.05</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.11</td>
<td>0.32</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.69</td>
<td>0.92</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.61</td>
<td>0.64</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.48</td>
<td>1</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.63</td>
<td>1.63</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.44</td>
<td>1.27</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.62</td>
<td>0.3</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.2</td>
<td>0.09</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.65</td>
<td>0.18</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.08</td>
<td>0.46</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.78</td>
<td>12.26</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.83</td>
<td>4.58</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.76</td>
<td>13.07</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.17</td>
<td>11.43</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.65</td>
<td>2.61</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td><strong>Forest productivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLP</td>
<td>All</td>
<td>0.63</td>
<td>1.63</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.44</td>
<td>1.27</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.62</td>
<td>0.3</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.2</td>
<td>0.09</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.65</td>
<td>0.18</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.08</td>
<td>0.46</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.78</td>
<td>12.26</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Bombuscaro</td>
<td>0.83</td>
<td>4.58</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.76</td>
<td>13.07</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.17</td>
<td>11.43</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.65</td>
<td>2.61</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td><strong>Canopy traits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaf toughness</td>
<td>Bombuscaro</td>
<td>0.2</td>
<td>0.09</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.65</td>
<td>0.18</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.08</td>
<td>0.46</td>
<td>27</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.78</td>
<td>12.26</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>SLA</td>
<td>Bombuscaro</td>
<td>0.83</td>
<td>4.58</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.76</td>
<td>13.07</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.17</td>
<td>11.43</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.65</td>
<td>2.61</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Foliar N</td>
<td>Bombuscaro</td>
<td>0.35</td>
<td>1.69</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.79</td>
<td>1.91</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cajanuma</td>
<td>0.27</td>
<td>1.60</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.65</td>
<td>0.17</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>Foliar P</td>
<td>Bombuscaro</td>
<td>0.85</td>
<td>0.08</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>San Francisco</td>
<td>0.80</td>
<td>0.14</td>
<td>14</td>
<td>2</td>
</tr>
</tbody>
</table>
|                  | Cajanuma       | 0.37        | 0.11     | 23            | 2         

Between 8% and 37% of the variance, 20-85% of the variance was explained in Bombuscaro, and 69-80% of the variance was explained in San Francisco.

4.3.2 Predictor importance

Regarding predictor importance, the predictors’ significance and coefficient magnitude varied slightly among the variables of forest biomass, productivity and canopy traits (Figure 4.3). In fact, elevation was an important predictor in all
models. We found a high correlation of elevation with the RGVI (r = 0.90, p<0.001) and EVI (r = -0.79, p<0.001) but not with RBVI. RBVI exhibited high importance in the models of FLP and NPPa and in the models of foliar P. Slope was only important in the model of NPPa and ’entropy’ statistics of the RGVI were only significant in the model of foliar P.

We further compared the best subset model with models fitted by only topographical, only spectral and only textural predictor variables to evaluate their predictive power (Appendix C Figure C.3). Regarding single predictor sets, AGBi and SLA were best explained by topographical metrics and AGB by textural metrics, whereas spectral metrics best explained all remaining response variables. Overall the best subset models (Table 4.2; Figure 4.3) outperformed the single predictor sets in all investigated models. In particular, for FLP, NPPa, foliar N and foliar P, the best subset model explained a serious amount of variance compared to the single-predictor sets.

### 4.3.3 Prediction maps

For all models, along the whole elevation gradient, we plotted the spatial variation of predicted...
Figure 4.4: Prediction maps of aboveground biomass (AGB, a), annual wood production (AGB\textsubscript{i}, b), annual fine litter production (FLP, c), aboveground net primary production (NPP\textsubscript{a}, d) and community-weighted means (CWM) of leaf toughness (e), the specific leaf area (SLA, f), foliar N (g) and foliar P concentration (h). Non-forest areas and no-data values were masked out using a Landsat-based classification of 2001 (gray = non-forest; Göttlicher et al., 2009). Elevation in m a.s.l. was added as contour lines for 1000 m, 2000 m and 3000 m.
response variables within forests (Figure 4.4). Visually, the prediction maps revealed a strong spatial congruence in most parts of the forest. In particular, the maps for AGB, AGB and SLA and the maps for FLP, NPP, foliar N and foliar P exhibited high similarities. For AGB, AGB and the SLA, a shift from high to low values was detected between the study site Bombuscaro and the other two study sites, San Francisco and Cajanuma, with the highest biomass, productivity rates and SLA values at elevations between 1000 m and 2000 m a.s.l. In contrast, on average, the annual FLP and NPP exhibited the highest productivity rates within San Francisco. The estimated mean annual NPP accounted for 4.79 ± 1.13 Mg ha\(^{-1}\) y\(^{-1}\) in the study area (Figure 4.4d). The three study sites (boxes in Figure 4.4d) exhibited variations in annual NPP, with an average of 5.11 ± 0.90 Mg ha\(^{-1}\) y\(^{-1}\) in Bombuscaro, 5.16 ± 1.23 Mg ha\(^{-1}\) y\(^{-1}\) in San Francisco, and 4.95 ± 1.18 Mg ha\(^{-1}\) y\(^{-1}\) in Cajanuma.

4.4 Discussion

Our study investigated topographical, multispectral and textural metrics derived from remote sensing data to model tropical forest biomass, productivity and canopy traits along a prominent elevation gradient. Models of biomass, productivity and canopy traits were derived with moderate-to-high predictive power along the whole elevation gradient, in addition to FLP and canopy traits within certain individual study sites. Although elevation was an important predictor in all models, we demonstrated that spectral and textural metrics explained additional variance, in particular, of biochemical canopy traits and related fine litter production.

4.4.1 Forest biomass and productivity

In our study area, forest biomass and productivity are driven by abiotic changes along the elevational gradient. Abiotic factors such as temperature, nutrient availability and soil moisture, decrease with increasing elevation, and therefore biomass and productivity decrease, also (Homeier et al., 2010; Malhi et al., 2004; Raich et al., 2006; Ryan et al., 2006). A post hoc analysis controlling for effects of elevation, slope and topographic position revealed that aboveground biomass and the annual wood production were mainly determined by elevation (results not shown). Models at single study sites therefore hardly explained variation in aboveground biomass and the annual wood production. As proposed by Wallis et al. (2017), tree species composition changes with increasing elevation. Therefore, we assume that AGB and AGB were mainly driven by compositional changes in trees among the three study sites, which can be retrieved from elevation.

For AGB, however, textural metrics were also important predictors. In agreement with Lu and Batistella (2005), we found moderate correlations between AGB and image textures (results not shown), and in accordance with Dube and Mutanga (2015b), we also identified that models fitted by multiple texture metrics slightly exceeded the predictive power of models fitted by spectral or topographical metrics only (Appendix C Figure C.3). We therefore suggest that image textures were able to discriminate between differences in canopy structure among
the three study sites and within each study site. Thus, biomass is more strongly related to canopy structure than to the initial topographical metrics.

In contrast to AGB, the model of FLP exhibited higher predictive power, and even the models within Bombuscaro and San Francisco still explained 61% and 48% of the variance, respectively. Models controlling for effects of topographical metrics revealed that spectral metrics explained additional variance in the residuals for FLP ($R^2=0.41$) and for NPP ($R^2=0.24$; results not shown). The litterfall, which is mostly derived from canopy leaves, is most likely more related to canopy characteristics such as the spectral response and its texture than is AGB. In contrast, AGB might be less related to the spectral signal and canopy structure than FLP; therefore, its model might be enhanced by the inclusion of environmental predictors not retrieved, such as soil conditions and water availability (Unger et al., 2012).

### 4.4.2 Canopy traits

Regarding the canopy trait models, our findings supported our hypothesis that multispectral satellite data are able to model the investigated canopy traits along the elevational gradient and within single study sites. However, there is an ongoing debate whether there is a direct or an indirect relationship between spectral reflectance data and biochemical leaf traits (Knyazikhin et al., 2013a; Lepine et al., 2016; Ollinger, 2011; Townsend et al., 2013). On the one hand, our results are in line with previous studies showing that the characteristics of multispectral wavelength bands and indices such as the near-infrared (NIR), the soil adjusted ratio vegetation index or the EVI are relevant to predict morphological traits such as the SLA at the canopy level across forests (Ali et al., 2017; Asner and Martin, 2008; Lymburner et al., 2000) or biochemical canopy traits, such as the foliar N (Lepine et al., 2016). On the other hand Knyazikhin et al. (2013b) and Knyazikhin et al. (2013a) proposed that foliar N cannot be estimated from the NIR without accounting for the canopy structure. A different opinion is that the relationship between NIR and foliar N might be derived via the correlation of foliar N with structural traits (Ollinger et al., 2013; Townsend et al., 2013). In response to Ollinger (2011), using coarser sensor data has been explored for modeling foliar N and total chlorophyll density (Lepine et al., 2016; Shiklomanov et al., 2016). Lepine et al. (2016) compared data from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) with data derived from Landsat TM data to model foliar N concentration. Although the Landsat-derived regression models were weaker than those from AVIRIS, the difference vegetation index still explained 61% of foliar N (Lepine et al., 2016). Shiklomanov et al. (2016) stated that the performance of Landsat 8 data for modeling certain leaf parameters is superior to that of Landsat 5 data, indicating that even slight differences in the wavelength locations of certain bands can be important for the retrieval of canopy traits.

Notwithstanding, regarding our models, elevation and topographic variability were prominent factors influencing the compositions of trees and their spectral and textural response (Wallis et al., 2017). As proposed by Asner and Martin (2016), elevational changes in canopy
traits are mediated by shifts in the canopy tree composition, and environmental filtering is mostly driven by soil fertility, light availability and air temperature (Asner and Martin, 2016; Long et al., 2011). The elevational and topographical gradients codetermine the distribution of tree height, for instance, with lower trees at higher elevation and ridges, which, in turn, are related to lower SLA values and higher leaf toughness (Jucker et al., 2018; Madani et al., 2018). Higher SLA values, in contrast, are found at lower elevations where water and nutrient availability are higher (Long et al., 2011). Although these environmental filters could be retrieved by elevation and the TPI, we stress that spectral metrics outperformed these topographical metrics in all canopy trait models and in the FLP model. The change in forest structure with increasing elevation is associated with lower chlorophyll content at higher altitudes, which in turn reduces the reflectance of green light and NIR and leads therefore to lower EVI and RGVI values. In contrast, the RBVI, which was not correlated with elevation, was an important predictor for biochemical canopy traits and the related annual FLP.

Controlling for topographical effects showed that spectral predictors explained 24% and 23% of the residuals from the models of the morphological canopy traits leaf toughness and SLA, respectively, and 33% and 44% of the residuals from the models of CWM foliar N and foliar P concentrations (results not shown). In contrast, textural metrics explained only 12% for the model of SLA and less than 10% of the residuals from the remaining canopy traits. Thus, spectral metrics are able to explain additional information in particular in biochemical canopy traits and related canopy production not retrieved from topographical information. Consequently, the predictive power of these models improved when topographical, spectral and textural metrics were used together as predictors.

Foliar N and foliar P, which are known as key traits controlling gross primary productivity, were both strongly correlated with FLP and NPP in our study area (Appendix C Figure C.4; Hättenschwiler et al., 2011; LeBauer and Treseder, 2008; Mercado et al., 2011). In contrast to leaf toughness and SLA, they exhibited a higher association with spectral metrics. In accordance with Smith et al. (2002) we therefore suggest that biochemical canopy traits can act as important surrogates of forest productivity and can be retrieved from remote sensing metrics.

4.4.3 Limitations

In a similar manner as for most MODIS and Landsat based tropical AGB models, our study might suffer from mixed pixels and a mismatch between pixel-size and field-plot area. Regarding mixed pixels, we tested for the study site San Francisco whether the proportional forest cover per pixel derived from a linear spectral unmixing had an effect on the predictive power of the models (results not shown; Göttlicher et al., 2009). However, the mixed pixel effect (forest – non-forest) was not significant in any model; therefore, we suggest that the effect of mixed pixels can be neglected in our study area. Regarding the mismatch between pixel size and field-plot area, the permanent forest plots are situated in homogenous natural forest stands, and their surroundings should not differ signifi-
cantly from the sampled in situ data. Only for AGB, which is neither a proportional measurement nor an average value, might the models be biased due to the mismatch in observed values per 20 m × 20 m plot and predicted values within the 30 m × 30 m pixels.

4.4.4 Conclusion and implications

This study investigated the relationships between topographical, spectral and textural remote sensing metrics, forest biomass, productivity and canopy traits. We showed that a moderate-to-high amount of variation in tropical forest productivity and in canopy traits can be explained by multispectral data at moderate resolution along an elevation gradient from 1000 m up to 3000 m a.s.l. We identified convergent trends in the spatial distribution of multiple canopy traits along the elevational gradient. We suggest that the investigated canopy traits were driven by similar ecological factors that change along the complex topography of our study area, which were best explained by a combination of topographical, spectral and textural metrics. In particular, spectral metrics were highly associated with FLP, NPP, and biochemical canopy traits. Considering future research, Sentinel-2 data will improve the retrieval of carbon-related variables in tropical montane forest due to their finer spectral and spatial resolution compared to Landsat images (Drusch et al., 2012; Malenovský et al., 2012; Pandit et al., 2018). The enhanced red-edge spectra might be able to more precisely model canopy traits (Houborg et al., 2015) and most likely also measures of forest productivity.

Acknowledgements

This study was conducted in the framework of the German Research Foundation (DFG) Research Unit FOR 816 (LE762/10, HO3296/2), and PAK 823-825 (subprojects A1 HO3296/4, C2 BE1780/34-1, BR1293/11, FA925/7-1, Z1698/8-1). We are grateful to the DFG (PAK 825/1) for funding our project, to the Ecuadorian Ministry of the Environment (MAE) for permission to conduct research, and to the foundation Nature and Culture International (NCI) for logistical support. We thank three anonymous reviewers for helpful comments on the manuscript in its initial form. Special thanks go to Roman Link, Greta Weitmann, Stefanie Lorenz, Philipp Obst and Nixon Cumbicus for field assistance (forest inventories and leaf trait sampling).

References

Asner, G., Martin, R., 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. Remote Sensing of En-


Safari, A., Sohrabi, H., 2016. Ability of Landsat-8 OLI derived texture metrics in estimating aboveground carbon stocks
of coppice oak forests. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLI-B8, 751–754. doi:10.5194/isprsarchives-XLI-B8-751-2016.


Spracklen, D.V., Righelato, R., 2014. Trop-
Chapter 4  Modeling forest biomass, productivity and canopy traits with multispectral data


Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., Yu, S., 2016. Examining Spectral Reflectance Sat-
References

Zvoleff, A., 2015. Glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R.

Zvoleff, A., 2015. Glcm: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R.
Chapter 5

Synthesis
5.1 Summary and conclusion

The development of spatially explicit indicators is of particular interest for biodiversity monitoring purposes. This is especially the case in remote and topographically complex regions such as tropical mountain rainforests. The opportunities to derive comprehensive environmental proxies from multispectral data for modeling biodiversity are manifold.

In this thesis, I focused on the potential of spectral and textural metrics for modeling patterns of tropical montane forest biodiversity in the Andes of Southern Ecuador. I hypothesized that models of biodiversity benefit from the inclusion of these metrics (H1) but that their predictive performances differ among the biodiversity variables, diversity indices and taxa studied (H2). My objectives were to (a) compare the performance of spectral and textural metrics in modeling taxonomic and functional diversity, different diversity indices and taxa, (b) assess the benefit of spectral and textural metrics as complementary predictors to topographical metrics, and (c) compare airborne and spaceborne multispectral sensors in modeling biodiversity. In three studies (Chapter 2-4) I have demonstrated that the performance of topographical, spectral, and textural metrics differ among biodiversity variables, investigated taxa and diversity indices studied (Figure 5.1). Consequently, H1 was only valid for certain investigated biodiversity variables which supported H2. This will be discussed as follows.

The success of spectral and textural metrics in modeling biodiversity variables mainly depended on the following factors: i) the biological entity or attribute, the diversity index and taxon under investigation; ii) the specific spectral or textural metrics used including the window size applied in texture calculation, closely related to iii) the spatial resolution of the sensor.

(i) The potential of multispectral remote sensing is closely related to the abiotic conditions driving the abundance and diversity of biological entities and attributes. As proposed by Peters et al. (2016), there is no general model that explains elevational species diversity of multiple taxa separately. Taxonomic diversity seems to depend on the taxon-specific resource requirements and their specific adaption to the environment. In line with this suggestion, I found varying predictor importance among topographical, spectral and textural metrics (Figure 5.1).

The elevational lapse in temperature influences the abundance and species richness of certain ectotherms (e.g., ants and pyraloid moths; Chapter 3). However, topography reveals also variation of abiotic factors such as nutrient supply and hydraulic conditions driving the composition of trees and their structure (Chapter 3; Jucker et al., 2018). Along with topographic features such as mountains and topographic positions, microclimate changes, which in turn influences the composition of trees and growth of vegetation (Chapter 3-4). According to the changing habitat conditions, the species composition of most taxonomic groups changed with increasing elevation (Chapter 3).
5.1 Summary and conclusion

Figure 5.1: Schematic overview of the benefit of spectral and textural metrics in modeling biodiversity variables. The potential of spectral metrics is defined as the complementary information explained to topographical metrics and the potential of textural metrics is defined as the complementary information explained to combined topographical and spectral metrics. The three corners represent the three sets of remotely sensed predictors surrogating changing abiotic conditions along with elevation, vegetation productivity and habitat structure. Symbols characterize measures of taxonomic and functional diversity. Symbols of taxa represent their species richness (for birds Shannon diversity). The position inside the triangle indicates the relative predictor importances; the center indicates that approx. 33% of the biodiversity variable are explained by each predictor set; dashed lines are the the 50% mark of the predictor set in the opposite corner. Bird species richness, for example, is highly determined by habitat structure surrogated by image textural metrics derived from high-spatial-resolution data but not by spectral or topographical metrics. No models were derived for arctiinae moths richness using high spatial resolution data and for bird species richness using moderate resolution data. Species turnover is not illustrated for clarity reasons. In general, a change in species composition across all taxa was found with elevation, as well as with a combination of spectral and textural metrics indicating that community compositions also change along a habitat structural gradient. Model results using topographical metrics only, topographical and spectral together, and combined topographical, spectral and textural metrics can be found in Appendix B Table B.6, Appendix D Table D.2, Table D.3 and Table D.4. AGB, aboveground biomass; AGBi, annual wood production; NPPa, aboveground net primary production; FLP, fine litter production; SLA, specific leaf area.
In contrast, other taxonomic groups have developed distinct features to cope with the change of abiotic conditions along with elevation. Some of the investigated biological entities and attributes were related to variations of habitat productivity or habitat structure (Chapter 2-4). The three studies demonstrated that spectral metrics, especially a broadband index combining the blue and red wavelength bands, have a high potential for modeling the richness of geometrid moth species (Appendix B Table B.6), in addition to biochemical canopy traits and related canopy production (Appendix B Table B.6, Appendix D Table D.2; Chapter 3-4). A high potential of textural metrics was found for the taxonomic diversity of certain taxa and biochemical canopy traits and their production (Chapter 4, Appendix D Figure D.1). Vegetation productivity and habitat structure thus drive ecosystem processes and functions of the canopy, which might be related to the species richness of these taxonomic groups, namely, birds (Chapter 2-3), trees, and geometrid and arctiinae moths (Chapter 3).

My findings also revealed a compositional change in species of most investigated taxa along the habitat structural gradient (Chapter 3). Considering all investigated biodiversity variables (viz. species richness and species turnover of trees, moths, birds and ants for the models at high spatial resolution and additional forest biomass, productivity and canopy traits for the models at moderate resolution), the summed proportion explained by textural metrics was about 40% of variance (complementary to topographic and spectral metrics) in diversity (Appendix D Table D.3, Table D.4). Whereas, topographic metrics explained about 50% of variance. This confirmed prior studies suggesting that habitat structure is an important driver of taxonomic and functional diversity (e.g., Bae et al., 2018), and emphasized the potential of textural metrics as remotely sensed proxies to model tropical montane biodiversity.

(ii) Regarding the predictor importance in all three studies, two points appeared to be important. On the one hand, in contrast to the EVI and NDVI, the broad-band combination of the blue and red wavelength bands was independent from elevation, and therefore, explained complementary information to topographical metrics. Consequently, this spectral index was responsible for the improved models of biochemical canopy traits (Chapter 4) using moderate spatial resolution data and probably also for the model of geometrid moth species richness (Chapter 3). On the other hand, the success of textural metrics depended on the moving window size. Generally, the models of species richness of birds and trees, and models of biochemical canopy traits and fine litter production benefitted from textural metrics calculated from multiple moving window sizes (Chapter 3, Appendix D Figure D.1). This is caused by the activity size (home range) of the investigated taxa that needs to be matched by the size of the moving window in texture calculation (Mairota et al., 2015). Multiple taxa simply require textural metrics derived from multiple window sizes as proxies for their home ranges. This fact was crucial for the models of functional diversity such as biochemical canopy traits and production. The window size in texture calculation, however,
is affected by the spatial resolution of multispectral data (see iii). This highlights that texture calculation should be performed carefully considering the choice of window sizes as well as the spatial resolution to match the appropriate scale of habitat structure.

(iii) The benefit of spectral and textural metrics varied not only among biological entities and attributes studied, but also among the spatial resolution of multispectral data. I investigated different scales using multispectral data at very high spatial resolution (2.5 m, Chapter 2), at high spatial resolution (0.3 m resampled to 6 m, Chapter 3) and at moderate spatial resolution (30 m, Chapter 4). Although the Landsat 8 metrics had the lowest spatial resolution, the predictability of most diversity measures was still high or even higher than in models using high spatial resolution data (Chapter 3, Appendix B Table B.6, Appendix D Table D.2). Therefore, I suggest that remote sensing data at moderate spatial resolution is still valuable because the spectral bands of Landsat-8 data are well defined.

However, it turned out that the spatial resolution could be a crucial factor determining the predictive power of the models and the benefit of spectral and textural metrics. For instance, in contrast to textural metrics, spectral metrics at moderate spatial resolution were not able to model the richness of geometrid moth species. I suggest that the spectral information was too broad and therefore, additional information about the habitat provided by textural metrics was necessary. Regarding bird species diversity, models at moderate resolution performed poor (Appendix D Table D.2). This result was most likely related to the small extent of bird species records. Consequently, the weak performance of the model highlights the need of operational sensors with high to very high spatial and spectral resolutions to model biodiversity at finer scales and small spatial extents.

To the best of my knowledge, this thesis provides the first evaluation of the potential of multispectral data for modeling tropical montane biodiversity considering taxonomic and functional diversity, multiple taxa and vegetation indices. As long as hyperspectral and Lidar data are restricted to airborne surveys and commercial spaceborne missions, multispectral remote sensing is still of great importance for monitoring biodiversity. Together, the findings of my research have important implications for the development of a remotely sensed indicator system of biodiversity. In sum, the following points are relevant:

- Biodiversity models could be enhanced through the use of specific broadband combinations using visible wavelength bands and multispectral image textures.
- The benefit of spectral and textural metrics depends on the biological entity or attribute studied.
- Multispectral remote sensing data at moderate resolution was able to model most biodiversity
variables, but for some taxa more complex predictors in form of multiple texture metrics were necessary (c.f. geometrid and arctiinae moths).

- The sampling extent of in situ data should match to the spatial resolution of remote sensing data (c.f. bird species richness)

Nonprofit organizations, that depend on open-source data archives such as those provided by the Landsat mission, benefit from the development of spatially explicit biodiversity indicators using multispectral data. Regarding the Aichi goals, especially target 11 and 14, monitoring in general and the identification of prioritization areas in particular, requires robust and spatially explicit indicators of biodiversity. The enhancement of biodiversity models using operational remote sensing data is, therefore, of value and can help to derive more accurate maps on biodiversity. Accordingly, the demonstrated potential of multispectral remote sensing for operational biodiversity modeling, contributes to both the scientific community and nonprofit organizations.

### 5.2 Perspectives

The presented findings of this thesis point to further investigations. These arise on different levels. Amongst others, to further enhance remote sensing driven models of biodiversity at a landscape scale, investigations can be related to either sensor improvements, choice of biodiversity indicators or statistical methods.

Sensor improvements have been intensively discussed in the last years by the Committee on Earth Observation Satellites (CEOS) owing to the need for operational satellites with sensors at higher spectral and spatial resolutions or with active sensors on-board (Kuenzer et al., 2014; Turner et al., 2013; Nagendra et al., 2013). This need has been addressed by a range of upcoming or recently launched new Earth observation satellites which provide free and open data at high spatial and spectral resolution. The Sentinel mission, of the Copernicus program, from the European Space Agency (ESA) is one of them. Sentinel-1, for example, includes four planned satellites with a synthetic aperture Radar (SAR) on-board. Two of these four satellites were already launched and first evaluations showed that Sentinel-1 will have a high potential in modeling vegetation structure (Schmidt et al., 2017). The ability to derive the three-dimensional structure of vegetation from space seems promising to model biomass and wood production directly. As shown in Chapter 2, enhanced delineation of vertical structure might also improve models of phylodiversity. Indicator derivation of functional leaf traits, in contrast, might profit from hyperspectral sensors at high spatial resolutions. In particular the models of biochemical canopy traits, investigated in Chapter 4, benefitted from spectral metrics and I suggest they would further benefit from higher spectral resolution data. This need might be answered by the Environmental Mapping and Analysis Program (EnMAP) that plans to launch a satellite with a hyperspectral sensor on-board and thus provide free hyperspectral data at moderate spatial resolution (30 m × 30 m) and high
temporal resolution (4 days revisit). But as suggested from previous studies and supported by my findings, carefully selected wavelength bands perform sometimes even better in modeling and mapping canopy traits and productivity, as well as taxonomic species diversity, than a high number of narrow wavelength bands (Shiklomanov et al., 2016). Therefore for certain research even a low number of well-defined wavelength bands might be appropriate. Promising, the Sentinel-2 mission will enhance recent land cover mapping possibilities. Sentinel-2 which has two satellites in orbit carries multispectral sensors at high spatial resolution (10-60 m) with a revisiting time of approx. 5 days (Malenovský et al., 2012). This will enable a more precise mapping of vegetation by leaf-chlorophyll content and vegetation indices (Berger et al., 2012). In combination with a textural approach Sentinel-2 data might further enhance models of taxonomic and functional diversity as horizontal canopy structure would be better depicted by textural metrics at higher spatial and spectral resolutions. In addition a higher spatial resolution would avoid independency problems that might arise from multiple plots within a moving window in texture calculation (see Appendix D Figure D.1). Up to now no cloud-free scene has been derived from Sentinel-2 for the study area. But it seems worthwhile to compute a composite from cloud-free pixels for the study area in order to test its benefits for biodiversity purposes in comparison to the Landsat-8 data assessed in Chapter 4.

With regard to biodiversity monitoring, a major challenge seems to be the combination of the obtained spatial indicators to represent biodiversity as a whole. Even if the proposed essential biodiversity variables can be tracked from space, their combination and weighting will be difficult. To account for the functioning of the ecosystem directly, plant functional types could be investigated. According to similarities in their spectral response, remote sensing metrics should be able to detect groups of plant functional types (Bonan et al., 2002). On the one hand they therefore provide finer scales than biomes. On the other hand they comprise many ecosystem processes at once. This enables a monitoring of plant functional types at larger scales and at the same time accounting for within-habitat variations. In addition, they may be used to study the effects of climate change on plant communities with similar functions (Bonan et al., 2002). With regard to this thesis, plant functional types could be, for example, derived from the results in Chapter 4. As plant functional types could be investigated at the community-level, the scale of analysis would match conservation decisions and, therefore, provide effective indicators for stakeholders.

The used partial least squares regression approach, the algorithm was able to handle a low number of observations (samples) but with constraints regarding the predictor importance. If multi-collinearity is present, no precise information on predictor importance were possible. If predictors are removed or added, the latent structure would change and therefore no precise inference on predictor influence can be obtained. In addition, there might be a variety of models with different predictor combinations that are statistically reasonable but could lead to different conclusions about the relationship between in situ data and remote sensing predictors. For this
reason, in this thesis, the indirect benefit of predictor sets has been analyzed. But for more precise information on predictor importance and robustness, further investigations are necessary. One statistical solution for assessing the importance of predictors in the given predictor sets in general (and not just in one model) would be sub-window permutation analysis, which addresses the problem of the variety of predictor combinations. Subsets of samples and predictors are used in N Monte-Carlo simulations (Li et al., 2010). This allows calculation of a conditional p-value for each variable without considering the influence of the remaining variables. However, this approach is rarely implemented nowadays in partial least squares regressions. I believe that a remote sensing based indicator system would highly benefit from this method since it specifies which indicators are the most informative and robust irrespective of additional predictors used.

In conclusion, various aspects can be addressed to increase the predictive power of biodiversity models and the robustness of their remotely sensed environmental predictors. These investigations will further emphasize the potential of recent and future satellite missions providing operational remote sensing data at low cost for monitoring biodiversity.

References


Zusammenfassung
Zusammenfassung


In dieser Arbeit wurde das Potenzial multispektraler Fernerkundungsdaten analysiert, taxonomische und funktionelle Aspekte der Biodiversität in einem tropischen Bergregenwald im Süden Ecuadors zu modellieren. Im Speziellen wurden Vegetationsindizes aus multispektralen Reflektanzen sowie auf diesen basierende Texturmaße verwendet. Dazu wurden (i) verschiedene taxonomische Gruppen und Diversitätsmaße (z.B. alpha-/beta-Diversität), untersucht, (ii) ein Vergleich zu topographischen Metriken gezogen, und (iii) Sensordaten mit hoher sowie moderater räumlicher Auflösung berücksichtigt.


Der relative Anteil an der erklärenden Varianz der Texturmetriken als zusätzliche Prädiktoren zu topographischen und spektralen Metriken betrug über alle Biodiversitätsmodelle hinweg
Acknowledgments
This work lasted more than four years, several stays in Ecuador, two published scientific articles, one in revision, two book chapters and the birth of my daughter. All this would not be possible without the contribution of several people.

First of all, I would like to thank my supervisors. Jörg Bendix gave me the freedom to deepen my work. His feedback, reviews and instructions helped me to find my way into science. I will always remember a brief but intense “career discussion” in the kitchen with him, where he gave me confidence in myself. Roland Brandl guided me through statistical problems, R-code and biological terminology. His statistical advice (best on the phone at 6 am) helped me a lot and I appreciate how much time he has invested, even at weekends. As Roland Brandl, Nina Farwig also gave me insights into ecology. Nina guided me and my colleague Yve during our field work in 2014. I thank her for many fruitful discussions, feedback and helpful comments.

Looking back on my time in Ecuador, I remember my first stay with many rainy days. I am glad that I was not alone in the forest, but accompanied by Yvonne Tiede and Jan Schlautmann. Thank you for the shared Oreo cookies and the relaxed days in Vilcabamba after many weeks full of work. I particularly enjoyed my second field stay in Ecuador. Not only because it rained less but also because I had the most enthusiastic and motivated field assistants with me. Thank you Leonie and Julian, I am glad that we are still in contact and I wish you all the best for your future. Not to forget Isabel and Annika and many other student assistants who also participated immensely in the collection of field data.

My work has benefitted from collaborating with many people. I would like to highlight two of my co-authors who have made a special effort to introduce me to their field of research. Konrad Fiedler and Jürgen Homeier kindly drew my attention to the importance of taxonomic knowledge and showed me their passion for insects and trees, respectively. I recently also came across some mails from Ingo Grass, who helped me very quickly with my phylogenetic analysis of birds during my stay in Ecuador. I would like to thank him (once again) for his comprehensive help.

The LCRS team is also worth mentioning. Although I worked most often from abroad, I always felt very comfortable in the working group. Sometimes I would have preferred to live “full time” in Marburg. My special thanks go to Birgit Kühne-Bialozyt, Katja Trachte and Boris Thies, who were good contacts when it came to organizational questions or for a short conversation in the corridor.

My thanks also go to Matthias Koschorreck, Benjamin Bleyhl and Philipp Keller for their more or less spontaneous and in the case of Philipp on-going proofreading and helpful comments. And of course to my parents for their support during my education.

At the beginning of this thesis I had a great backing from my roommates Benni and Saskia. In the end, however, it was Philipp’s support that helped me to finish this work. He was not only my partner, but also my friend, my reviewer and a peer. I am very grateful to share my life with you and Elsa.
Appendix A

Chapter 2
Table A.1: Explained variance of the first four principal components of the PCA of bird abundance matrix. Each principal component represents the distribution of bird species along its ordination axis. We chose the first principal component as a proxy for a typical bird assemblage in our study area because it explains most variance in community composition.

<table>
<thead>
<tr>
<th>Importance of components:</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation</td>
<td>2.3285</td>
<td>1.5448</td>
<td>1.32811</td>
<td>1.29228</td>
</tr>
<tr>
<td>Proportion of Variance</td>
<td>0.2074</td>
<td>0.09126</td>
<td>0.06746</td>
<td>0.06387</td>
</tr>
<tr>
<td>Cumulative Proportion</td>
<td>0.2074</td>
<td>0.29862</td>
<td>0.36608</td>
<td>0.42994</td>
</tr>
</tbody>
</table>

Table A.2: Loadings of first principal component based on the covariance matrix.

<table>
<thead>
<tr>
<th>Species</th>
<th>Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanocorax yncas</td>
<td>0.261</td>
</tr>
<tr>
<td>Myioborus miniatus</td>
<td>0.232</td>
</tr>
<tr>
<td>Henicorhina leucophrys</td>
<td>0.216</td>
</tr>
<tr>
<td>Basileuterus tristriatus</td>
<td>0.212</td>
</tr>
<tr>
<td>Pyrrhomyias cinnamomeus</td>
<td>0.196</td>
</tr>
<tr>
<td>Scytalopus micropterus</td>
<td>0.183</td>
</tr>
<tr>
<td>Lepidocolaptes lacrymiger</td>
<td>0.182</td>
</tr>
<tr>
<td>Anisognathus somptuosus</td>
<td>0.180</td>
</tr>
<tr>
<td>Basileuterus coronatus</td>
<td>0.174</td>
</tr>
<tr>
<td>Synallaxis azarae</td>
<td>0.164</td>
</tr>
<tr>
<td>Aulacorhynchus prasinus</td>
<td>0.160</td>
</tr>
<tr>
<td>Colibri thalassinus</td>
<td>0.139</td>
</tr>
<tr>
<td>Penelope barbata</td>
<td>0.132</td>
</tr>
<tr>
<td>Diglossa albilatera</td>
<td>0.132</td>
</tr>
<tr>
<td>Creurgops verticalis</td>
<td>0.123</td>
</tr>
<tr>
<td>Momotus aequatorialis</td>
<td>0.122</td>
</tr>
<tr>
<td>Thraupis cyanocephala</td>
<td>0.120</td>
</tr>
<tr>
<td>Ochthoea cinnamomeiventris</td>
<td>0.116</td>
</tr>
<tr>
<td>Chlorospingus parvirostris</td>
<td>0.108</td>
</tr>
<tr>
<td>Zimmerius chrysops</td>
<td>0.107</td>
</tr>
<tr>
<td>Boissonneaua matthewsii</td>
<td>0.106</td>
</tr>
<tr>
<td>Rupicola peruvianus</td>
<td>0.104</td>
</tr>
<tr>
<td>Ocreatus underwoodii</td>
<td>0.102</td>
</tr>
<tr>
<td>Iridosornis analis</td>
<td>0.100</td>
</tr>
<tr>
<td>Tangara xanthocephala</td>
<td>0.095</td>
</tr>
<tr>
<td>Chlorospingus flavigularis</td>
<td>0.090</td>
</tr>
<tr>
<td>Species</td>
<td>Frequency</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Syndactyla subalaris</td>
<td>0.086</td>
</tr>
<tr>
<td>Myadestes ralloides</td>
<td>0.085</td>
</tr>
<tr>
<td>Heliodoxa leadbeateri</td>
<td>0.084</td>
</tr>
<tr>
<td>Chlorothraupis frenata</td>
<td>0.082</td>
</tr>
<tr>
<td>Adelomyia melanogenys</td>
<td>0.081</td>
</tr>
<tr>
<td>Aglaiocercus kingi</td>
<td>0.081</td>
</tr>
<tr>
<td>Dendroica fusca</td>
<td>0.077</td>
</tr>
<tr>
<td>Phaethornis syrmatophorus</td>
<td>0.077</td>
</tr>
<tr>
<td>Mionectes striaticollis</td>
<td>0.068</td>
</tr>
<tr>
<td>Chamaepetes goudottii</td>
<td>0.062</td>
</tr>
<tr>
<td>Myiodynastes chrysocephalus</td>
<td>0.059</td>
</tr>
<tr>
<td>Tangara nigroviiridis</td>
<td>0.056</td>
</tr>
<tr>
<td>Doryfera ludovicae</td>
<td>0.050</td>
</tr>
<tr>
<td>Heliodoxa rubinoides</td>
<td>0.048</td>
</tr>
<tr>
<td>Thraupis episcopus</td>
<td>0.048</td>
</tr>
<tr>
<td>Tyrannus melancholicus</td>
<td>0.048</td>
</tr>
<tr>
<td>Catharus ustulatus</td>
<td>0.044</td>
</tr>
<tr>
<td>Tangara labradorides</td>
<td>0.043</td>
</tr>
<tr>
<td>Wilsonia canadensis</td>
<td>0.043</td>
</tr>
<tr>
<td>Troglodytes solstitialis</td>
<td>0.042</td>
</tr>
<tr>
<td>Zonotrichia capensis</td>
<td>0.040</td>
</tr>
<tr>
<td>Margarornis squamiger</td>
<td>0.039</td>
</tr>
<tr>
<td>Haplophaedia aureliae</td>
<td>0.038</td>
</tr>
<tr>
<td>Grallaria hypoleuca</td>
<td>0.036</td>
</tr>
<tr>
<td>Geotrygon frenata</td>
<td>0.035</td>
</tr>
<tr>
<td>Coeligena coeligena</td>
<td>0.033</td>
</tr>
<tr>
<td>Thryothorus euophrys</td>
<td>0.031</td>
</tr>
<tr>
<td>Eutoxeres aquila</td>
<td>0.025</td>
</tr>
<tr>
<td>Myiobius villosus</td>
<td>0.025</td>
</tr>
<tr>
<td>Pseudotriscus pelzelni</td>
<td>0.025</td>
</tr>
<tr>
<td>Piaya cayana</td>
<td>0.024</td>
</tr>
<tr>
<td>Nothocercus bonapartei</td>
<td>0.024</td>
</tr>
<tr>
<td>Premnoplex brunnescens</td>
<td>0.023</td>
</tr>
<tr>
<td>Odontophorus speciosus</td>
<td>0.023</td>
</tr>
<tr>
<td>Poecilotriccus ruficeps</td>
<td>0.022</td>
</tr>
<tr>
<td>Urosticte ruficrissa</td>
<td>0.022</td>
</tr>
<tr>
<td>Species</td>
<td>Percentage</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Arremon torquatus</td>
<td>0.021</td>
</tr>
<tr>
<td>Vireo leucophrys</td>
<td>0.021</td>
</tr>
<tr>
<td>Dendrocincla tyrannina</td>
<td>0.019</td>
</tr>
<tr>
<td>Diglossa sittoides</td>
<td>0.019</td>
</tr>
<tr>
<td>Leptopogon rufipictus</td>
<td>0.019</td>
</tr>
<tr>
<td>Xenops rutilans</td>
<td>0.019</td>
</tr>
<tr>
<td>Elaenia albiceps</td>
<td>0.018</td>
</tr>
<tr>
<td>Rhynchoscyclus fulvipectus</td>
<td>0.018</td>
</tr>
<tr>
<td>Nyctibius griseus</td>
<td>0.015</td>
</tr>
<tr>
<td>Contopus fumigatus</td>
<td>0.015</td>
</tr>
<tr>
<td>Grallaria squamigera</td>
<td>0.013</td>
</tr>
<tr>
<td>Phyllomyias cinereiceps</td>
<td>0.012</td>
</tr>
<tr>
<td>Glaucidium jardinii</td>
<td>0.009</td>
</tr>
<tr>
<td>Turdus fuscater</td>
<td>0.008</td>
</tr>
<tr>
<td>Siptornis striaticollis</td>
<td>0.007</td>
</tr>
<tr>
<td>Colaptes rubiginosus</td>
<td>0.006</td>
</tr>
<tr>
<td>Coeligena torquata</td>
<td>0.005</td>
</tr>
<tr>
<td>Mecocerculus minor</td>
<td>0.003</td>
</tr>
<tr>
<td>Megascops petersoni</td>
<td>0.002</td>
</tr>
<tr>
<td>Colibri coruscans</td>
<td>0.001</td>
</tr>
<tr>
<td>Premnornis guttuligera</td>
<td>0.001</td>
</tr>
<tr>
<td>Chalcostigma ruficeps</td>
<td>0.001</td>
</tr>
<tr>
<td>Hemispingus frontalis</td>
<td>0.001</td>
</tr>
<tr>
<td>Parula pitiayumi</td>
<td>0.000</td>
</tr>
<tr>
<td>Xiphorhynchus triangularis</td>
<td>-0.002</td>
</tr>
<tr>
<td>Grallaria rufula</td>
<td>-0.003</td>
</tr>
<tr>
<td>Xiphocolaptes promeropirhynchus</td>
<td>-0.004</td>
</tr>
<tr>
<td>Lesbia nuna</td>
<td>-0.005</td>
</tr>
<tr>
<td>Myiarchus cephalotes</td>
<td>-0.007</td>
</tr>
<tr>
<td>Anairetes parulus</td>
<td>-0.008</td>
</tr>
<tr>
<td>Cistothorus platensis</td>
<td>-0.008</td>
</tr>
<tr>
<td>Elaenia pallatangae</td>
<td>-0.008</td>
</tr>
<tr>
<td>Lafresnaya lafresnayi</td>
<td>-0.008</td>
</tr>
<tr>
<td>Scytalopus parkeri</td>
<td>-0.009</td>
</tr>
<tr>
<td>Myiophobus flavicans</td>
<td>-0.010</td>
</tr>
<tr>
<td>Cyclarhis gujanensis</td>
<td>-0.012</td>
</tr>
<tr>
<td>Species</td>
<td>Score</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Phyllomyias nigrocapillus</td>
<td>-0.012</td>
</tr>
<tr>
<td>Pipreola riefferii</td>
<td>-0.012</td>
</tr>
<tr>
<td>Pharomachrus auriceps</td>
<td>-0.015</td>
</tr>
<tr>
<td>Diglossa cyanea</td>
<td>-0.017</td>
</tr>
<tr>
<td>Arremon brunneinucha</td>
<td>-0.018</td>
</tr>
<tr>
<td>Tangara parzudakii</td>
<td>-0.019</td>
</tr>
<tr>
<td>Turdus fulviventris</td>
<td>-0.020</td>
</tr>
<tr>
<td>Metallura tyrianthina</td>
<td>-0.020</td>
</tr>
<tr>
<td>Tangara vassorii</td>
<td>-0.021</td>
</tr>
<tr>
<td>Campephilus pollens</td>
<td>-0.021</td>
</tr>
<tr>
<td>Contopus sordidulus</td>
<td>-0.021</td>
</tr>
<tr>
<td>Myiophobus lintoni</td>
<td>-0.021</td>
</tr>
<tr>
<td>Ochthoeca diadema</td>
<td>-0.021</td>
</tr>
<tr>
<td>Coeligena lutetiae</td>
<td>-0.022</td>
</tr>
<tr>
<td>Mecocerculus stictopterus</td>
<td>-0.022</td>
</tr>
<tr>
<td>Thamnophilus unicolor</td>
<td>-0.023</td>
</tr>
<tr>
<td>Diglossa humeralis</td>
<td>-0.024</td>
</tr>
<tr>
<td>Veniliornis dignus</td>
<td>-0.025</td>
</tr>
<tr>
<td>Saltator cinctus</td>
<td>-0.025</td>
</tr>
<tr>
<td>Colaptes rivolii</td>
<td>-0.026</td>
</tr>
<tr>
<td>Piranga rubriceps</td>
<td>-0.030</td>
</tr>
<tr>
<td>Atlapetes latinuchus</td>
<td>-0.032</td>
</tr>
<tr>
<td>Basileuterus luteoviridis</td>
<td>-0.033</td>
</tr>
<tr>
<td>Myiarchus tuberculifer</td>
<td>-0.036</td>
</tr>
<tr>
<td>Accipiter ventralis</td>
<td>-0.039</td>
</tr>
<tr>
<td>Trogon personatus</td>
<td>-0.040</td>
</tr>
<tr>
<td>Lipaugus fuscocinereus</td>
<td>-0.045</td>
</tr>
<tr>
<td>Turdus serranus</td>
<td>-0.045</td>
</tr>
<tr>
<td>Cinnycerthia unirufa</td>
<td>-0.048</td>
</tr>
<tr>
<td>Haplospiza rustica</td>
<td>-0.052</td>
</tr>
<tr>
<td>Sericossypha albocristata</td>
<td>-0.053</td>
</tr>
<tr>
<td>Cinnycerthia olivascens</td>
<td>-0.054</td>
</tr>
<tr>
<td>Grallaria nuchalis</td>
<td>-0.055</td>
</tr>
<tr>
<td>Hemitriccus granadensis</td>
<td>-0.058</td>
</tr>
<tr>
<td>Myornis senilis</td>
<td>-0.061</td>
</tr>
<tr>
<td>Anisognathus lacrymosus</td>
<td>-0.064</td>
</tr>
<tr>
<td>Species</td>
<td>Value</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Myioborus melanocephalus</td>
<td>-0.067</td>
</tr>
<tr>
<td>Eriocnemis vestita</td>
<td>-0.073</td>
</tr>
<tr>
<td>Diglossa caerulescens</td>
<td>-0.076</td>
</tr>
<tr>
<td>Grallaricula nana</td>
<td>-0.086</td>
</tr>
<tr>
<td>Chlorornis riefferii</td>
<td>-0.102</td>
</tr>
<tr>
<td>Buthraupis montana</td>
<td>-0.104</td>
</tr>
<tr>
<td>Pseudocolaptes boissonneautii</td>
<td>-0.108</td>
</tr>
<tr>
<td>Patagioenas fasciata</td>
<td>-0.137</td>
</tr>
<tr>
<td>Heliangelus amethysticollis</td>
<td>-0.145</td>
</tr>
<tr>
<td>Pionus senilis</td>
<td>-0.157</td>
</tr>
<tr>
<td>Synallaxis unirufa</td>
<td>-0.170</td>
</tr>
<tr>
<td>Scytalopus unicolor</td>
<td>-0.186</td>
</tr>
<tr>
<td>Chlorospingus ophthalmicus</td>
<td>-0.197</td>
</tr>
</tbody>
</table>
Table A.3: Excerpt of species loaded with highest and lowest values in the first principal component (based on covariance matrix). Associations between loading and traits were not found. Abbreviations are as follows. Foraging strata: U = understory, M = midstory, C = canopy; Center of abundance: UT = upper tropical, UM = upper montane, HT = hill tropical, MM = middle montane; Relative abundance: C = common, F = fairly common, U = uncommon, P = patchily distributed; Min/Max elevation: L = lowlands; Habitats: F4 = mountain evergreen forest, F15 = secondary forest, N3 = semihumid/humid montane scrub, F1 = tropical lowland evergreen forest, F11 = pine-oak forest, F7 = tropical deciduous forest, E = edge; Conservation priority: 1 = urgent, 2 = high, 3 = medium; Research priority: 1 = high, 2 = medium, 3 = low. *species were not listed in Stotz et al. (1996).

<table>
<thead>
<tr>
<th>Species</th>
<th>PC1 Loadings</th>
<th>Sensivity</th>
<th>Foraging strata</th>
<th>Center of abundance</th>
<th>Relative abundance</th>
<th>Min</th>
<th>Max</th>
<th>Habitats</th>
<th>Conservation priority</th>
<th>Research priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanocorax yncas</td>
<td>0.2608</td>
<td>Low</td>
<td>C</td>
<td>UT</td>
<td>F/P</td>
<td>L</td>
<td>2800</td>
<td>F4,F7,F15</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Myioborus miniatus</td>
<td>0.2323</td>
<td>Low</td>
<td>M/C</td>
<td>UT</td>
<td>C</td>
<td>600</td>
<td>2500</td>
<td>F4,F15,F11</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Henicorhina leucophrys</td>
<td>0.2162</td>
<td>Medium</td>
<td>U</td>
<td>UT</td>
<td>C</td>
<td>900</td>
<td>3000</td>
<td>F4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Basileuterus tristriatus</td>
<td>0.2123</td>
<td>Medium</td>
<td>U</td>
<td>UT</td>
<td>C</td>
<td>800</td>
<td>500</td>
<td>F4,F15</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Pyrrhomyias cinnamomeus</td>
<td>0.1963</td>
<td>Medium</td>
<td>C</td>
<td>UT/MM</td>
<td>C</td>
<td>1000</td>
<td>3350</td>
<td>F4,F4E</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Scytalopus micropterus*</td>
<td>0.1834</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lepidocolaptes lacrymiger*</td>
<td>0.1824</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Anisognathus somptuosus</td>
<td>0.1798</td>
<td>Medium</td>
<td>M/C</td>
<td>MM</td>
<td>C</td>
<td>900</td>
<td>2300</td>
<td>F4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Basileuterus coronatus</td>
<td>0.1738</td>
<td>Medium</td>
<td>U</td>
<td>MM</td>
<td>C</td>
<td>1400</td>
<td>2800</td>
<td>F4,F15</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Chlorornis riefferii</td>
<td>-0.1015</td>
<td>Medium</td>
<td>C</td>
<td>MM</td>
<td>F</td>
<td>300</td>
<td>350</td>
<td>F4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Buthraupis montana</td>
<td>-0.1038</td>
<td>Medium</td>
<td>C</td>
<td>MM</td>
<td>C</td>
<td>2000</td>
<td>500</td>
<td>F4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Pseudocolaptes boissonneautii</td>
<td>-0.1079</td>
<td>Medium</td>
<td>M/C</td>
<td>MM</td>
<td>F</td>
<td>1400</td>
<td>400</td>
<td>F4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Patagioenas fasciata</td>
<td>-0.1374</td>
<td>Medium</td>
<td>C</td>
<td>UM</td>
<td>F</td>
<td>900</td>
<td>600</td>
<td>F4,F11,F15</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Heliangulus amethysticollis</td>
<td>-0.1448</td>
<td>Medium</td>
<td>U/M</td>
<td>UM</td>
<td>F</td>
<td>800</td>
<td>300</td>
<td>F4,F15</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Pionus senilis</td>
<td>-0.1566</td>
<td>Medium</td>
<td>C</td>
<td>HT</td>
<td>U</td>
<td>L</td>
<td>1600</td>
<td>F4,F1,F15</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Synallaxis unirufa</td>
<td>-0.1701</td>
<td>Medium</td>
<td>U</td>
<td>UM</td>
<td>?</td>
<td>1700</td>
<td>300</td>
<td>F4,F5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Scytalopus unicolor</td>
<td>-0.1860</td>
<td>High</td>
<td>U</td>
<td>UM</td>
<td>?</td>
<td>2000</td>
<td>3150</td>
<td>N3?</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Chlorospingus ophthalmicus</td>
<td>-0.1966</td>
<td>Medium</td>
<td>U/M</td>
<td>UT</td>
<td>C</td>
<td>1000</td>
<td>2500</td>
<td>F4,F15</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>
Table A.4: Predictors for each diversity measure selected owing to filtering in backward selection. For each predictor the base layer of the texture approach, the texture algorithm, and the window size applied for calculating the image texture is given. Window sizes define the surrounding of each pixel in which image textures were calculated. Given window sizes refer to the pixel size of the sensor. * Both window sizes were selected.

<table>
<thead>
<tr>
<th>Shannon diversity</th>
<th>Phylodiversity</th>
<th>Bird community</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIR CO/DI/EN/HO 45×45</td>
<td>NDVI CC/CO 45×45</td>
<td>NIR ME/VA **</td>
</tr>
<tr>
<td>PRI SM 45×45</td>
<td>PRI EN 45×45</td>
<td>NIR CC/CO/DI/EN 45×45</td>
</tr>
<tr>
<td>PRI CC/EN/SM 45×45</td>
<td>PRI EN 45×45</td>
<td>PRI CC 45×45</td>
</tr>
<tr>
<td>HOME ME/VA 113×113</td>
<td>DEM VA 113×113</td>
<td>DEM ME <strong>/VA</strong></td>
</tr>
<tr>
<td>CH ME 13×13/VA 113×113</td>
<td>MH ME/VA 13×13</td>
<td>MH ME/VA 113×113</td>
</tr>
<tr>
<td>VDR ME 113×113</td>
<td>SLOPE ME/VA 113×113</td>
<td>E51 ME/VA 113×113</td>
</tr>
<tr>
<td>LiDAR (first-order texture)</td>
<td>E55 ME 13×13/VA**</td>
<td>E58 ME/VA 13×13</td>
</tr>
<tr>
<td></td>
<td>E57 ME **/VA 113×113</td>
<td>E68 ME/VA 13×13</td>
</tr>
<tr>
<td></td>
<td>E68 ME 13×13</td>
<td></td>
</tr>
</tbody>
</table>
Figure A.1: Forest loss between 2001-2014 for the study area derived from Landsat forest classifications (Hansen et al., 2013); classification data has been downloaded as “year of gross forest cover loss event” for the granule with top-left corner at 0N, 80W (Source: Hansen/UMD/Google/USGS/NASA). Rectangular buffers relate to the extent of the textural approach of optical and Lidar metrics. Forest loss within the Reserva Biológica San Francisco and the Podocarpus National Park was probably caused by landslide events which are very common in this region. Forest loss next to non-forest areas was probably caused by slash-and-burn events. Only one bird point count was affected by forest loss in 2011. Since Lidar data was acquired in 2012 and this point was not an outlier within the Lidar models, we conclude that habitat change had only a minor impact on our Lidar models.
Figure A.2: Figure A2 Scatterplot matrix of diversity proxies. Correlations are based on the Pearson correlation moment.
Figure A.3: Figure A.3. Response plots of all six models fitted by either Lidar or optical texture metrics. A leave-one-out validation was performed.
Benefit of multi-sensor models

Other research areas, such as carbon stock estimations and forest stand inventories, also benefit from the combined use of different sensor data (Baccini et al., 2012; Kellndorfer et al., 2010; Saatchi et al., 2008; Walker et al., 2007). Some studies have explored the beneficial use of coupling image textures with elevation or habitat type (St-Louis et al., 2009, 2006), but to our knowledge, no study to date has explored the benefit of multi-sensor models fitted by combined Lidar and texture metrics to address different aspects of species diversity. Elevation metrics in addition to optical-based metrics might be beneficial for modeling bird species richness (Sheeren et al., 2014). Consequently, we tested for all bird diversity proxies (a) the benefit of models fitted by textural image metrics and complex Lidar metrics and (b) the benefit of models fitted with elevation and slope in addition to texture metrics of optical images. We only found a benefit among bird community models, where Lidar and texture metrics together yielded a LOO $R^2$ of 0.85 and texture models with elevation and slope in addition yielded a LOO $R^2$ of 0.81. All other models did not perform better than the best single-sensor models based on a comparison of LOO $R^2$ and RMSE values.


Appendix B

Chapter 3
Table B.1: Number of rarefied and extrapolated samples of tree and moth species covering 70% of the asymptotic richness. Sum of each row reveal the total number of samples.

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>No. of samples using rarefied species richness</th>
<th>No. of samples using extrapolated species richness</th>
<th>No. of samples using observed species richness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>28</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Geometridae</td>
<td>20</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>18</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Table B.2: Stress values for the first two dimensions of NMDS ordinations.

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>1 Dimension</th>
<th>2 Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>0.136</td>
<td>0.08</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.101</td>
<td>0.071</td>
</tr>
<tr>
<td>Geometridae</td>
<td>0.098</td>
<td>0.057</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.14</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table B.3: Texture statistics derived from gray-level co-occurrence matrices calculated from optical metrics (see Table 1). For the moving window approach that shifts in all directions, one window size was considered within the core analysis (3 pixels × 3 pixels). Image texture metrics derived from an additional window size (17 pixels × 17 pixels) to cope with different habitat scales of all taxonomic groups was used for a second model (Table B.5).

Texture statistics  

**Mean**  
\[ ME = \sum_{i,j=0}^{N-1} (p_{ij}) \]

**Entropy**  
\[ EN = \sum_{i,j=0}^{N-1} p_{ij} (-\ln p_{ij}) \]

**Correlation**  
\[ CC = \sum_{i,j=0}^{N-1} p_{ij} \left( \frac{(i-ME)(j-ME)}{\sqrt{VA_iVA_j}} \right) \]

*With \( p_{ij} = \frac{V_{ij}}{\sum_{i,j=0}^{N-1} V_{ij}} \), where \( V_{ij} \) is the value in cell i, j and N is the number of rows or columns.*
Table B.4: Pearson correlation coefficients of habitat indicators and tree growth parameters. For all derived habitat indicators, we performed a correlation analysis with tree growth parameters derived from a plot-based inventory (Homeier et al., unpublished). Pearson’s correlation coefficients depended on different numbers of samples. Data on stem diameter at breast height, above ground biomass increment (AGB), and a measure of aboveground net primary production (NPP$_a$; combination of AGB and leaf litter production) were collected, and tree species were sampled ($n = 50$). In addition, we measured the C/N ratio of leaves in understory plants and sampled ant species ($n = 27$). Significant correlation coefficients are highlighted in bold. DEM, Digital elevation model; SLOPE, slope in degrees; TPI, topographical position index; NIR, near-infrared band; NDVI, normalized difference vegetation index; ARI, anthocyanin reflectance index.

<table>
<thead>
<tr>
<th>Habitat indicators</th>
<th>Stem size</th>
<th>AGB$_i$</th>
<th>NPP$_a$</th>
<th>C/N ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEM</td>
<td>0.38*</td>
<td>-0.64*</td>
<td>-0.64*</td>
<td>0.41</td>
</tr>
<tr>
<td>SLOPE</td>
<td>0.28</td>
<td>0.13</td>
<td>0.29</td>
<td>-0.46</td>
</tr>
<tr>
<td>TPI</td>
<td>-0.20</td>
<td>0.28</td>
<td>0.13</td>
<td>0.41</td>
</tr>
<tr>
<td>NIR mean</td>
<td>-0.32</td>
<td>0.59*</td>
<td>0.78*</td>
<td>-0.08</td>
</tr>
<tr>
<td>NDVI mean</td>
<td>-0.40*</td>
<td>0.62*</td>
<td>0.75*</td>
<td>-0.19</td>
</tr>
<tr>
<td>ARI mean</td>
<td>0.11</td>
<td>0.09</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td>NIR correlation</td>
<td>-0.22</td>
<td>0.19</td>
<td>0.27</td>
<td>-0.31</td>
</tr>
<tr>
<td>NDVI correlation</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.11</td>
<td>-0.44</td>
</tr>
<tr>
<td>ARI correlation</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.39</td>
</tr>
<tr>
<td>NIR entropy</td>
<td>-0.36</td>
<td>0.58*</td>
<td>0.74*</td>
<td>0.19</td>
</tr>
<tr>
<td>NDVI entropy</td>
<td>-0.49*</td>
<td>0.64*</td>
<td>0.56*</td>
<td>-0.42</td>
</tr>
<tr>
<td>ARI entropy</td>
<td>-0.32</td>
<td>0.35</td>
<td>0.13</td>
<td>-0.55*</td>
</tr>
</tbody>
</table>

*p < 0.01
Table B.5: Results of partial-least square regressions with implemented backward selection and leave-one-out cross validation (LOO) using additional 'mean' and 'entropy' texture statistics derived from a second moving window of 3 pixels × 3 pixels (18 m × 18 m), which resulted in 18 predictors in total. All resulting models included texture statistics from the additional smaller window after backward selection. LOO $R^2$ values higher than those in Table 4 are in bold.

<table>
<thead>
<tr>
<th>Species diversity measures</th>
<th>Taxonomic group</th>
<th>$R^2$</th>
<th>LOO $R^2$</th>
<th>No. of latent vectors</th>
<th>No. of predictors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Species richness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td></td>
<td>0.49</td>
<td><strong>0.22</strong></td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td></td>
<td>0.77</td>
<td>0.48</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Geometridae</td>
<td></td>
<td>0.75</td>
<td>0.57</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Arctiinae</td>
<td></td>
<td>0.38</td>
<td><strong>0.09</strong></td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Ants</td>
<td></td>
<td>0.8</td>
<td>0.74</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Birds</td>
<td></td>
<td>0.61</td>
<td><strong>0.27</strong></td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td><strong>NMDS I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td></td>
<td>0.99</td>
<td>0.98</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td></td>
<td>0.94</td>
<td>0.87</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Geometridae</td>
<td></td>
<td>0.96</td>
<td><strong>0.93</strong></td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Arctiinae</td>
<td></td>
<td>0.96</td>
<td><strong>0.93</strong></td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Ants</td>
<td></td>
<td>0.94</td>
<td>0.90</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Birds</td>
<td></td>
<td>0.80</td>
<td><strong>0.68</strong></td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td><strong>NMDS II</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td></td>
<td>0.42</td>
<td>0.29</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td></td>
<td>0.89</td>
<td><strong>0.82</strong></td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Geometridae</td>
<td></td>
<td>0.79</td>
<td>0.62</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Arctiinae</td>
<td></td>
<td>0.83</td>
<td><strong>0.72</strong></td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Ants</td>
<td></td>
<td>0.82</td>
<td><strong>0.75</strong></td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Birds</td>
<td></td>
<td>0.66</td>
<td><strong>0.53</strong></td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>
Table B.6: Results of partial least-squares regressions predicting taxa diversities using (1) only topographical metrics (DEM, SLOPE, and TPI) as predictors and (2) both topographical and spectral (NIR, NDVI, ARI) metrics. Because of the low dimension of predictors, we omitted the backward selection and chose two latent vectors for all models. The LOO (pseudo-) $R^2$ values are calculated as $1 - \text{SSE}/\text{SST}$, where SST is the corrected total sum of squares of the response, and SSE is the sum of squared errors for cross-validated predictions. Negative LOO $R^2$ values indicate an inability of the models (on average) to explain any of the variability in investigated taxa diversity among sample plots. DEM, Digital elevation model; SLOPE, slope in degrees; TPI, topographical position index; NIR, near-infrared band; NDVI, normalized difference vegetation index; ARI, anthocyanin reflectance index.

<table>
<thead>
<tr>
<th>Species diversity</th>
<th>Taxonomic group</th>
<th>Topographical metrics only</th>
<th>Topographical and spectral metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>LOO $R^2$</td>
<td>$R^2$</td>
</tr>
<tr>
<td>Species richness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td>0.20</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.82</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td>Geometridae</td>
<td>0.02</td>
<td>-0.68</td>
<td>0.64</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.11</td>
<td>-0.27</td>
<td>0.11</td>
</tr>
<tr>
<td>Ants</td>
<td>0.80</td>
<td>0.75</td>
<td>0.81</td>
</tr>
<tr>
<td>Birds</td>
<td>0.15</td>
<td>-0.01</td>
<td>0.16</td>
</tr>
<tr>
<td>NMDS I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.96</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Geometridae</td>
<td>0.96</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.93</td>
<td>0.86</td>
<td>0.95</td>
</tr>
<tr>
<td>Ants</td>
<td>0.21</td>
<td>-0.03</td>
<td>0.87</td>
</tr>
<tr>
<td>Birds</td>
<td>0.74</td>
<td>0.70</td>
<td>0.74</td>
</tr>
<tr>
<td>NMDS II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td>0.34</td>
<td>0.26</td>
<td>0.07</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.04</td>
<td>-0.87</td>
<td>0.52</td>
</tr>
<tr>
<td>Geometridae</td>
<td>0.01</td>
<td>-0.86</td>
<td>0.30</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.06</td>
<td>-0.86</td>
<td>0.41</td>
</tr>
<tr>
<td>Ants</td>
<td>0.61</td>
<td>0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>Birds</td>
<td>0.01</td>
<td>-0.20</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Appendix C

Chapter 4
**Table C.1:** Summary statistics of forest biomass, productivity variables and community weighted means (CWMs) of functional leaf traits among all sampled study sites (all) and the single study sites Bombuscaro at approximately 1000 m a.s.l., San Francisco at approximately 2000 m a.s.l. and Cajanuma at approximately 3000 m a.s.l. AGB, aboveground biomass; AGB$_i$, annual wood production; NPP$_a$, aboveground net primary production; FLP, fine litter production; SLA, specific leaf area.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Unit</th>
<th>Study site</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest biomass</td>
<td>AGB</td>
<td>All</td>
<td>29.83</td>
<td>143.3</td>
<td>423.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>84.27</td>
<td>182.52</td>
<td>338.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>29.83</td>
<td>158.01</td>
<td>423.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>45.04</td>
<td>89.36</td>
<td>182.7</td>
</tr>
<tr>
<td>Forest productivity</td>
<td>AGB$_i$</td>
<td>All</td>
<td>0.22</td>
<td>1.88</td>
<td>6.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>1.05</td>
<td>2.83</td>
<td>4.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>0.80</td>
<td>2.15</td>
<td>6.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>0.22</td>
<td>0.65</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>All</td>
<td>2.78</td>
<td>5.41</td>
<td>9.18</td>
</tr>
<tr>
<td></td>
<td>FLP</td>
<td>All</td>
<td>2.78</td>
<td>5.41</td>
<td>9.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>3.62</td>
<td>5.76</td>
<td>8.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>4.49</td>
<td>6.78</td>
<td>9.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>2.78</td>
<td>3.68</td>
<td>5.11</td>
</tr>
<tr>
<td></td>
<td>NPP$_a$</td>
<td>All</td>
<td>3.15</td>
<td>4.33</td>
<td>5.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>5.49</td>
<td>8.59</td>
<td>11.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>5.39</td>
<td>8.93</td>
<td>13.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>3.15</td>
<td>4.33</td>
<td>5.89</td>
</tr>
<tr>
<td>Leaf toughness</td>
<td>kN m$^{-1}$</td>
<td>All</td>
<td>0.51</td>
<td>1.23</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>0.66</td>
<td>0.9</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>0.51</td>
<td>1.07</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>0.99</td>
<td>1.72</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>SLA</td>
<td>All</td>
<td>32.19</td>
<td>74.93</td>
<td>149.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>76.53</td>
<td>95.59</td>
<td>110.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>47.76</td>
<td>79.85</td>
<td>149.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>32.19</td>
<td>49.33</td>
<td>76.92</td>
</tr>
<tr>
<td></td>
<td>Foliar N</td>
<td>All</td>
<td>9.04</td>
<td>16.07</td>
<td>27.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>14.21</td>
<td>18.29</td>
<td>21.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>12.93</td>
<td>18.61</td>
<td>27.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>9.04</td>
<td>11.32</td>
<td>14.69</td>
</tr>
<tr>
<td></td>
<td>Foliar P</td>
<td>All</td>
<td>0.32</td>
<td>0.73</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bombuscaro</td>
<td>0.56</td>
<td>0.76</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Francisco</td>
<td>0.55</td>
<td>0.95</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cajanuma</td>
<td>0.32</td>
<td>0.49</td>
<td>0.83</td>
</tr>
</tbody>
</table>
Table C.2: Classification of model quality with respect to the coefficient of determination ($R^2$).

<table>
<thead>
<tr>
<th>$R^2$ range</th>
<th>Model quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.2</td>
<td>Poor</td>
</tr>
<tr>
<td>0.2-0.4</td>
<td>Weak</td>
</tr>
<tr>
<td>0.4-0.6</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.6-0.8</td>
<td>Good</td>
</tr>
<tr>
<td>&gt;0.8</td>
<td>Strong</td>
</tr>
</tbody>
</table>
**Figure C.1:** Observed versus predicted values for PLSR models of forest biomass and productivity variables at individual study sites. First column corresponds to Figure 2. The dashed line represents the 1:1 line (y=x). Hollow points = calculated values, solid fill = predicted values. Aboveground biomass (AGB) in Mg ha$^{-1}$, and variables of forest productivity in Mg ha$^{-1}$ yr$^{-1}$. AGB$_i$, annual wood production; NPP$_a$, aboveground net primary production; FLP, fine litter production.
Figure C.2: Figure C2 Observed versus predicted values for PLSR models of canopy traits at individual study sites. First column corresponds to Figure 2. The dashed line represents the 1:1 line (y=x). Hollow points = calculated values, solid fill = predicted values. Community weighted means of leaf toughness in kN m$^{-1}$, of specific leaf area (SLA) in cm$^2$ g$^{-1}$ and of foliar N and P in mg g$^{-1}$. FLP, fine litter production.
Figure C.3: Comparison of LOO validated R² values for different model runs: predicting biomass, productivity and community weighted mean canopy traits using (1) the best subset of predictor variables (Figure 4.3), (2) only topographical metrics (DEM, SLOPE, and TPI) as predictors and (3) only spectral metrics (EVI, RGVI, RBVI) and (4) only textural metrics (EVI entropy, EVI correlation, RBVI entropy, RGVI entropy). Because of the low dimension of predictors, we omitted the backward selection for models fitted by topographical, spectral and textural metrics only and chose two latent vectors for their corresponding models. AGB, aboveground biomass; AGBᵢ, annual wood production; NPPᵢ, aboveground net primary production; FLP, fine litter production; SLA, specific leaf area.
Figure C.4: Figure C4 Scatterplot matrix of forest productivity variables and functional leaf traits. Correlations are based on the Pearson correlation moment. AGB, aboveground biomass; AGB\(_i\), annual wood production; NPP\(_a\), aboveground net primary production; FLP, fine litter production; SLA, specific leaf area
Appendix D

Additional information on the chapters 2-4
(not published)
Table D.1: From chapter 2 the hypothesis arose that the horizontal heterogeneity of vertical structure calculated as Lidar-derived texture metrics might improve predictions of bird diversity. For this reason, I calculated image textures using the same texture statistics and window sizes from chapter 2 using Lidar metrics (see Table 2.2, Table 2.3). This resulted in 188 Lidar-based texture metrics (Lidar texture). In addition, as a consequence of chapter 3 it remained unclear whether (textural) Lidar metrics are able to model the second bird community composition. Therefore, I derived models for the second principal component. As shown, Lidar-based models benefitted only slightly from the textural approach. In comparison with chapter 2, a higher predictive power was only found for phylo-α-diversity using Lidar-based texture metrics. Lidar models also exceeded multispectral texture metrics in modeling the second bird community composition. However, multispectral texture models were able to explain 39% in variation of the second community composition. Thus, it can be assumed that the second principal component indicates a change in species composition along a habitat structural gradient. The best model results are in bold.

<table>
<thead>
<tr>
<th>Diversity measure</th>
<th>Dataset</th>
<th>No. of predictors</th>
<th>No. of latent vectors</th>
<th>R²</th>
<th>RMSE</th>
<th>LOO R²</th>
<th>LOO RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon diversity</td>
<td>Lidar</td>
<td>5</td>
<td>2</td>
<td>0.26</td>
<td>0.85</td>
<td>0.07</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Lidar texture</td>
<td>7</td>
<td>2</td>
<td>0.26</td>
<td>0.85</td>
<td>0.03</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Multispectral texture</td>
<td>7</td>
<td>4</td>
<td>0.57</td>
<td>0.64</td>
<td><strong>0.42</strong></td>
<td>0.75</td>
</tr>
<tr>
<td>Phylo-α-diversity</td>
<td>Lidar</td>
<td>13</td>
<td>2</td>
<td>0.60</td>
<td>0.62</td>
<td>0.35</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Lidar texture</td>
<td>22</td>
<td>2</td>
<td>0.70</td>
<td>0.54</td>
<td><strong>0.52</strong></td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Multispectral texture</td>
<td>8</td>
<td>2</td>
<td>0.24</td>
<td>0.86</td>
<td>0.09</td>
<td>0.94</td>
</tr>
<tr>
<td>Bird community 1 (PC1)</td>
<td>Lidar</td>
<td>11</td>
<td>3</td>
<td>0.86</td>
<td>0.37</td>
<td><strong>0.81</strong></td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Lidar texture</td>
<td>7</td>
<td>3</td>
<td>0.85</td>
<td>0.38</td>
<td>0.80</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>Multispectral texture</td>
<td>8</td>
<td>2</td>
<td>0.85</td>
<td>0.38</td>
<td>0.80</td>
<td>0.44</td>
</tr>
<tr>
<td>Bird community 2 (PCA2)</td>
<td>Lidar</td>
<td>7</td>
<td>3</td>
<td>0.68</td>
<td>0.56</td>
<td><strong>0.58</strong></td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>Lidar texture</td>
<td>8</td>
<td>2</td>
<td>0.68</td>
<td>0.55</td>
<td>0.57</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Multispectral texture</td>
<td>9</td>
<td>3</td>
<td>0.56</td>
<td>0.65</td>
<td>0.39</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Table D.2: Results of partial-least squares regressions with implemented backward selection and leave-one-out (LOO) cross validation for species richness of trees, moths, ants and birds, the corresponding species turnover (by means of non-metric multidimensional scaling, NMDS), and forest biomass, its productivity and key canopy traits (measured as community weighted means) using a set of spatial predictors derived from the Landsat image used in chapter 4. As predictor metrics the same topographical, spectral and textural metrics were used as in chapter 4 but with an additional window size (19 pixels × 19 pixels) for texture calculations resulting in 17 predictors. See also Figure D1. The LOO (pseudo-) $R^2$ values are calculated as $1 - \text{SSE}/\text{SST}$, where SST is the corrected total sum of squares of the response, and SSE is the sum of squared errors for cross-validated predictions. Negative LOO $R^2$ values indicate an inability of the models (on average) to explain any of the variability in investigated taxa diversity among sample plots.

<table>
<thead>
<tr>
<th>Biodiversity variables</th>
<th>Taxa</th>
<th>Topographical metrics</th>
<th>Topographical and spectral metrics</th>
<th>Topographical, spectral, and textural metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species richness</td>
<td>Trees</td>
<td>0.12</td>
<td>0.19</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>0.58</td>
<td>0.51</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Geometridae</td>
<td>-0.64</td>
<td>-0.45</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>Arctiinae</td>
<td>-0.16</td>
<td>-0.82</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>0.73</td>
<td>0.66</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>-0.19</td>
<td>-0.45</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Trees</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>0.95</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Species turnover 1</td>
<td>Geometridae</td>
<td>0.93</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>(NMDS I)</td>
<td>Arctiinae</td>
<td>0.83</td>
<td>0.93</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>-0.12</td>
<td>-</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>-0.23</td>
<td>-</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>Trees</td>
<td>0.28</td>
<td>0.32</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Pyraloidea</td>
<td>-0.05</td>
<td>0.43</td>
<td>0.87</td>
</tr>
<tr>
<td>Species turnover 2</td>
<td>Geometridae</td>
<td>-0.50</td>
<td>0.19</td>
<td>0.64</td>
</tr>
<tr>
<td>(NMDS II)</td>
<td>Arctiinae</td>
<td>-0.38</td>
<td>0.15</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Ants</td>
<td>0.47</td>
<td>0.40</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Birds</td>
<td>-0.19</td>
<td>-0.19</td>
<td>0.35</td>
</tr>
<tr>
<td>Forest biomass</td>
<td>AGB</td>
<td>0.19</td>
<td>0.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Forest productivity</td>
<td>AGB$_i$</td>
<td>0.35</td>
<td>0.35</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>FLP</td>
<td>0.35</td>
<td>0.63</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>NPP$_a$</td>
<td>0.44</td>
<td>0.58</td>
<td>0.71</td>
</tr>
<tr>
<td>Canopy traits</td>
<td>Leaf toughness</td>
<td>0.48</td>
<td>0.60</td>
<td>0.63</td>
</tr>
<tr>
<td>(measured as community</td>
<td>SLA</td>
<td>0.67</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td>weighted mean)</td>
<td>Foliar N</td>
<td>0.41</td>
<td>0.61</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Foliar P</td>
<td>0.15</td>
<td>0.54</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Table D.3: Relative contributions of predictor sets at moderate spatial resolution based on the coefficients of determination in Table D.2. To account for correlations among the predictor sets, the contribution for the spectral predictor set has been calculated as the derived explained variance by topographical and spectral metrics minus the explained variance by topographical metrics only. The relative contributions are here used as an approximation to determine the benefit of the single predictor sets. In the last row the relative contributions were summed up (in %)

<table>
<thead>
<tr>
<th>Biodiversity variables</th>
<th>Relative contributions in explained biodiversity for each predictor set in %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Topographical metrics</td>
</tr>
<tr>
<td>Species richness</td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td>0.32</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.81</td>
</tr>
<tr>
<td>Geometridae</td>
<td>-</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>-</td>
</tr>
<tr>
<td>Ants</td>
<td>1.00</td>
</tr>
<tr>
<td>Birds</td>
<td>-</td>
</tr>
<tr>
<td>Trees</td>
<td>0.99</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>0.99</td>
</tr>
<tr>
<td>Geometridae</td>
<td>1.03</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.89</td>
</tr>
<tr>
<td>Ants</td>
<td>-</td>
</tr>
<tr>
<td>Birds</td>
<td>-</td>
</tr>
<tr>
<td>Trees</td>
<td>0.41</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>-</td>
</tr>
<tr>
<td>Geometridae</td>
<td>-</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>-</td>
</tr>
<tr>
<td>Ants</td>
<td>0.57</td>
</tr>
<tr>
<td>Birds</td>
<td>-</td>
</tr>
<tr>
<td>Forest biomass</td>
<td>AGB</td>
</tr>
<tr>
<td></td>
<td>AGB&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
<tr>
<td>Forest productivity</td>
<td>Litter</td>
</tr>
<tr>
<td></td>
<td>NPP&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>Toughness</td>
</tr>
<tr>
<td>Canopy traits</td>
<td>SLA</td>
</tr>
<tr>
<td></td>
<td>FoliarN</td>
</tr>
<tr>
<td></td>
<td>FoliarP</td>
</tr>
<tr>
<td>Relative contribution across all biodiversity variables</td>
<td>50</td>
</tr>
</tbody>
</table>
Table D.4: Relative contributions of predictor sets at high spatial resolution based on the coefficients of determination in Appendix B Table B5 and B6. To account for correlations among the predictor sets, the contribution for the spectral predictor set has been calculated as the derived explained variance by topographical and spectral metrics minus the explained variance by topographical metrics only. For the textural predictor set the explained variance by topographical and spectral (the higher one) has been subtracted. The relative contributions are here used as an approximation to determine the benefit of the single predictor sets. In the last row the relative contributions were summed up (in %).

<table>
<thead>
<tr>
<th>Biodiversity variables</th>
<th>Relative contributions in explained biodiversity for each predictor set in %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Topographical metrics</td>
</tr>
<tr>
<td>Species richness</td>
<td></td>
</tr>
<tr>
<td>Trees</td>
<td>0.36</td>
</tr>
<tr>
<td>Pyraloidea</td>
<td>1.00</td>
</tr>
<tr>
<td>Geometridae</td>
<td>0.89</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>1.00</td>
</tr>
<tr>
<td>Ants</td>
<td>1.00</td>
</tr>
<tr>
<td>Birds</td>
<td>-</td>
</tr>
<tr>
<td>NFDS I</td>
<td>0.92</td>
</tr>
<tr>
<td>Arctiinae</td>
<td>0.90</td>
</tr>
<tr>
<td>Ants</td>
<td>0.68</td>
</tr>
<tr>
<td>Birds</td>
<td>-</td>
</tr>
<tr>
<td>Relative contribution across all biodiversity variables</td>
<td>49</td>
</tr>
</tbody>
</table>
Figure D.1: In a prior version of the submitted manuscript in chapter 4, bigger window sizes were used in texture calculations which resulted in stronger model performances. Here the gain of explained variance measured as the coefficient of determination against the image textural metrics measured within different window sizes is shown. The corresponding textural metrics were subsequently added to a model including elevation, the topographical position index (TPI) and the RBVI as predictors. The RBVI was chosen as a spectral metric because it did not correlate with topography and therefore explained complementary information. To figure out which textural metrics would explain additional information in the eight response variables multiple window sizes were investigated. Overall, the gain in additional explained variance was marginal in the case of above-ground biomass, annual production of woody biomass (AGBi) and the morphological canopy traits leaf toughness and specific leaf area (SLA). For the remaining response variables textural metrics from big window sizes reached a higher gain in explained variance. However, I removed these predictor metrics in the submitted version of the article as the moving windows caused independency problems of the textural predictor metrics. As the impact of multiple plots within the same texture window could not be quantified, further studies with an enhanced sampling design are necessary to analyze the potential of image textures calculated at bigger window sizes in modelling forest biomass, its productivity and canopy traits. FLP = annual fine litter production, NPPa = net primary production. All canopy traits were measured as community-weighted means.
Erklärung
Hiermit versichere ich, dass ich meine Dissertation mit dem Titel

**Modeling tropical montane forest biodiversity**

–

**The potential of multispectral remote sensing**


Unterschrift (Christine Isabeau Bernarde Wallis)
Marburg an der Lahn, September 2018