TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

ENVIRONMENTAL VARIATION, NICHES OF INDICATOR SPECIES AND FOREST STRUCTURE IN AMAZONIA

Lassi Suominen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AII OSA – TOM. 399 | BIOLOGICA – GEOGRAPHICA – GEOLOGICA | TURKU 2023

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

ENVIRONMENTAL VARIATION, NICHES OF INDICATOR SPECIES AND FOREST STRUCTURE IN AMAZONIA

Lassi Suominen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. AII OSA – TOM. 399 | BIOLOGICA – GEOGRAPHICA – GEOLOGICA| TURKU 2023

University of Turku

Faculty of Science
Department of Biology
Doctoral programme in Biology, Geography and Geology

Supervised by

Adjunct Professor Kalle Ruokolainen
Departmenf of Biology
University of Turku
Turku, Finland

Professor Hanna Tuomisto
Department of Biology
University of Turku
Turku, Finland

Reviewed by

Doctor Alberto Vicentini
Instituto Nacional de Pesquisas da Amazônia (INPA)
Manaus, Brazil

Associate Professor David Zelený
Institute of Ecology and Evolutionary Biology
National Taiwan University, Taiwan

Opponent

Professor Jouko Rikkinen
University of Helsinki

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-9298-0 (PRINT)
ISBN 978-951-29-9299-7 (PDF)
ISSN 0082-6979 (Print)
ISSN 2343-3183 (Online)
Painosalama, Turku, Finland 2023

TURUN YLIOPISTO
Matemaattis-luonnontieteellinen tiedekunta
Biologian laitos
Ekologia
Lassi Suominen: Ympäristön vaihtelu, indikaattorilajien ekologiset lokerot ja metsän rakenne Amazoniassa
Väitöskirja, 157 s.
Biologian, maantieteen ja geologian tohtoriohjelma
Kesäkuu 2023

TIIVISTELMÄ

Amazonia on laajin jäljellä oleva trooppinen sademetsäalue. Sellaisena se on globaalisti tärkeä niin hiilen ja veden kiertokulkujen kuin luonnon monimuotoisuudenkin kannalta. Kuten useimmilla tropiikin alueilla, metsäkato on myös Amazoniassa huolestuttavan nopeaa. Sademetsiin kohdistuu monenlaisia maankäyttöpaineita, kuten ravinnontuotantoa, kaivostoimintaa ja puutavaran hankintaa. Vaikka sademetsäalue näyttää ilmakuvissa varsin yhtenäiseltä, alueen maaperä on huomattavan vaihtelevaa. Tämä vaihtelu puolestaan heijastuu eliölajiston vaihteluun. Myös metsän rakenteessa on alueellisia eroja. Maaperä- ja lajistovaihtelun kartoittaminen on nopeasti etenevän metsäkadon oloissa erityisen tärkeää, sillä maaperän ominaisuudet vaikuttavat tuottavuuteen – niin alkuperäisen metsäekosysteemin tuottavuuteen kuin ravinnontuotantoonkin. Maaperän ominaisuudet vaikuttavat myös siihen, minkä tyyppistä lajistoa alueella esiintyy. Maaperän suuri vaihtelevuus merkitsee usein myös suurta metsätyyppien vaihtelua ja siten alueellisesti korkeaa lajimäärää. Kartoitus on siten tärkeää myös lajiston suojelun kannalta. Kartoituksen hidasteena on ollut Amazonian laajuus ja monien alueiden vaikeapääsyisyys, suuri lajimäärä ja puutteellinen tieto alueen lajiston ekologiasta. Indikaattorilajien eli ilmaisinlajien käyttö on tehokas tapa nopeuttaa kartoitusta. Tässä väitöskirjassa käytän kahta suurta, viimeisten kolmen vuosikymmenen aikana kerättyä aineistoa Amazonian indikaattorilajeista. Ammennan Euroopan lehti- ja havumetsävyöhykkeiden pitkästä perinteestä, jossa indikaattorilajeja käytetään maaperän viljavuuden ja metsätyypin osoittamiseen. Tutkin, voitaisiinko samaa soveltaa Amazoniassa. Etsin potentiaalisia indikaattorilajeja mallintamalla niiden ekologisia lokeroita tärkeän maaperämuuttujan, emäskationikonsentraation, suhteen. Tutkin myös yleisiä kasviekologisia kysymyksiä, kuten sitä, ovatko yleiset ja runsaat lajit aina ekologisesti laaja-alaisia generalisteja. Näihin tarkoituksiin käytän HOF-malleja ja painotetun keskiarvon menetelmää. Testaan, kuinka tarkasti indikaattorilajeilla voi ennustaa erilaisia maaperän ominaisuuksia. Ennustamisessa sovellan *k*-lähin naapuri -menetelmää (*k*-NN) ja painotettua keskiarvokalibrointia. Tutkin myös maaperän vaikutusta alankosademetsän rakenteeseen. Suoritan maastomittauksia, joissa selvitän puiden tiheyttä eri kokoluokissa sekä latvuksen aukkoisuutta. Andien rinteiden pilvimetsissä tutkin metsän rakenteen ja mikroilmaston välistä yhteyttä, ja näiden vaikutusta epifyyttisten kasvien runsauteen ja monimuotoisuuteen. Suoritan maastossa metsän rakenteen ja epifyyttien runsauden mittauksia, ja automaattisten tallentimien avulla mikroilmaston seurantaa. Yhteensä viidessä osatyössä saan selville, että: 1) Sekä Melastomataceae -heimo että *Adiantum* ja *Lindsaea* -saniaissuvut sisältävät useita lupaavia indikaattorilajeja. Lisäksi kummankin ryhmän lajien optimit asettuvat emäskationigradientille siten, että mihin tahansa gradientin osaan tulee useita kapea-alaisia indikaattorilajeja, 2) Paikallisesti runsaat lajit eivät ole aina generalisteja, vaan ne voivat olla myös erikoistuneita tiettyyn maaperätyyppiin, 3) Etenkin emäskationikonsentraation ennustaminen indikaattorilajien avulla on varsin tarkkaa, mutta myös kaliumpitoisuutta on mahdollista ennustaa melko tarkasti. Muiden maaperämuuttujien suhteen menetelmä ei ole yhtä hyvä, 4) Metsän rakenteen erot eri maaperätyyppien välillä osoittautuvat vaikeiksi todentaa maastomittauksin, mutta eroja on myöhemmin löydetty kaukokartoitusmenetelmin, 5) Vuoristosademetsän olosuhteet muuttuvat epifyyteille epäsuotuisammiksi, kun latvuksen aukkoisuus lisääntyy. Tätä kehityskulkua vahvistavat sekä ilmastonmuutos että metsien hakkuut.

UNIVERSITY OF TURKU
Faculty of Science
Department of Biology
Ecology
LASSI SUOMINEN: Environmental variation, niches of indicator species and forest structure in Amazonia
Doctoral Dissertation, 157 pp.
Doctoral Programme in Biology, Geography and Geology
June 2023

ABSTRACT

Amazonia is the largest remaining tropical rainforest. As such, it is globally important both in terms of carbon and water cycles as well as biodiversity. As in most areas of the tropics, deforestation is also alarmingly fast in the Amazon. Rainforests are subject to many types of land use pressures, such as food production, mining and timber extraction. Although the rainforest area looks quite uniform in aerial photographs, the soil in the area is remarkably variable. This variation, in turn, is reflected in the variation of species composition. There are also regional differences in the structure of the forest. Mapping the variation in soils and species composition is particularly important in conditions of rapidly advancing deforestation, as soil properties affect productivity - both the productivity of the original forest ecosystem and food production. The properties of the soil also affect the assemblage of species present in the area. Large variability of the soil often also means a great variety of forest types and thus a regionally high number of species. Mapping is therefore also important for conservation purposes. The extent of Amazonia and the inaccessibility of many areas, the large number of species and the lack of knowledge about species ecology have slowed down the mapping. The use of indicator species is an effective way to speed up mapping. In this dissertation, I use two large datasets on Amazonian indicator species collected over the last three decades. I draw from the long tradition of Europe's deciduous and coniferous forest zones, where indicator species are used to indicate soil fertility and forest type. I am looking into whether the same could be applied in Amazonia. I look for potential indicator species by modelling their ecological niches in relation to an important soil variable, the concentration of base cations. I also study general plant ecological questions, such as whether common and abundant species are always ecologically wide-ranging generalists. For these purposes, I use HOF models and weighted averaging. I test how accurately different soil properties can be predicted with indicator species. In prediction, I apply the k-nearest neighbour method (*k*-NN) and weighted averaging calibration. I also study the influence soil has on the structure of the lowland rainforest understorey. I carry out field measurements of stem density in different size classes of trees, and of canopy openness. In the cloud forests of the Andean slopes, I study the relationship between the forest structure and microclimate, and their effect on the abundance and diversity of epiphytic plants. I carry out forest structure measurements and estimate the abundance of epiphytes in the field and record the microclimate with automatic data loggers. In five research papers, I find out that: 1) Both the Melastomataceae family and the fern genera *Adiantum* and *Lindsaea* contain several promising indicator species. In addition, the optima of the species in both groups are spread along the base cation gradient in such a way that several narrow-niche indicator species exist in all parts of the gradient, 3) Locally abundant species are not always ecological generalists, but can also be specialised to a certain soil type, 3) Soil base cation concentration can be accurately predicted using indicator, but it is also possible to predict the potassium content quite accurately. Regarding other soil variables, the method is not as good, 4) The differences in forest structure between different soil types prove to be difficult to verify with field measurements, but differences have later been found with remote sensing methods, 5) The conditions in the premontane cloud forest become less favourable for epiphytes with increasing canopy openness. This development is reinforced by both climate change and deforestation.

Table of Contents

List of Original Publications

This dissertation is based on the following original publications and manuscripts which will hereafter be referred to in the text by their Roman numerals.

I	Ruokolainen, K., Suominen, L. & Tuomisto, H. (2022) Distributions of Amazonian Melastomataceae species along a soil gradient. In: Goldenberg, R., Michelangeli, F.A. & Almeda, F. (eds.), *Advances in Melastomataceae Systematics and Biology*. Springer.
II	Tuomisto, H., Suominen, L., Alonso, A., Cárdenas, G., Lehtonen, S., Moulatlet, G.M., Peres, E., Sirén, A., Weigelt, P. & Zuquim. G. (2023) Species-soil relationships of Amazonian indicator plants and their consistency across regions. *Manuscript.*
III	Suominen, L., Ruokolainen, K., Tuomisto, H., Llerena, N. & Higgins, M.A. (2013) Predicting soil properties from floristic composition in western Amazonian rain forests: performance of k -nearest neighbour estimation and weighted averaging calibration. *Journal of Applied Ecology*, 50, 1441-1449.
IV	Suominen, L., Ruokolainen, K., Pitkänen, T. & Tuomisto, H. (2015) Similar understorey structure in spite of edaphic and floristic dissimilarity in Amazonian forests. *Acta Amazonica*, 45, 393-404.
V	Toivonen, J.M., Suominen, L., Gonzales-Inca, C.A., Trujillo Paucar, G. & Jones, M.M. (2017). Environmental drivers of vascular and non-vascular epiphyte abundance in tropical pre-montane cloud forests in Northern Peru. *Journal of Vegetation Science*, 28, 1198-1208.

The original publications have been reproduced with the permission of the copyright holders.

1 Introduction

Understanding what grows where, and why, is an important goal for ecologists and conservationists in Amazonia. The world's largest remaining tropical rainforest is facing threats from deforestation and climate change. Knowledge of plant species distributions, and of the environmental variables that influence them, is important for conservation planning and sustainable use of natural resources. Challenges are the low density of many rare plant species, which may be the most important for conservation, and the lack of fine-scale maps of environmental variation, such as soils. Indicator species can act as surrogates for both soils and species composition, and can therefore help in increasing the much-needed field data more quickly. In this thesis, I study indicator species, with a goal to advance the understanding of rainforest community ecology, and to produce results that are useful in land-use planning.

Environmental variables cause differences not only in species composition, but also in the structure of the vegetation. At the broadest scale, this is reflected in the tapestry of biomes that covers the Earth. Within biomes, such as the tropical rainforest, vegetation structure is affected by environmental variables, such as soil fertility and dry season length. The structure of the vegetation itself can also act as a factor affecting the performance of many plants. One way this works is through the shading that taller plants cast on lower ones. Such effects are perhaps strongest in epiphytic plant communities that literally grow on the vegetation itself and thereby have the availability of light, water and nutrients at least partly regulated by other plants. The role of the physical and chemical environment in controlling the structure of vegetation and, on the other hand, the role of such environmentally controlled features of vegetation in explaining the growth conditions of other plants is the second big theme in my thesis.

1.1 Environmental correlates of plant community composition in Amazonia

According to the dispersal assembly theory, variation in the relative abundances of plant species at any given locality depends mostly on the dispersal capabilities of the species and their relative abundances in the surrounding metacommunity (Hubbell

2001, Rosindell et al. 2011). In that theory, species are competitively equivalent within their trophic level. In contrast, the niche theory predicts that differences in species traits and niches have a major role in community assembly, and hence plant communities change in predictable ways along environmental gradients (Harpole & Tilman 2006, Kraft & Ackerly 2014). In niche assembly, the competitive fitness of species differs between habitats: it is not very common for a plant species to be simultaneously tolerant of different environmental extremes, e.g. drought and waterlogging (Silvertown 2004), although examples of this do exist, such as the vegetation in the seasonally flooded white-sand forests of Amazonia (Damasco et al. 2013).

If the variation in plant community composition in Amazonia could be explained by the dispersal assembly theory only, plant communities would show a universally decreasing similarity with increasing geographical distance. Such distance decay is indeed a phenomenon that characterises Amazonian tree communities at both regional and continental scales (Honorio Coronado et al. 2009). However, changes in plant species composition in Amazonia have also often been shown to be more strongly correlated with environmental similarity than with geographical distance (Phillips et al. 2003, Tuomisto et al. 2003 c, Kristiansen et al. 2012).

This relatively strong geological control of floristic composition is sometimes difficult to reconcile with the fact that there are a number of tree species that are able to dominate forests thousands of kilometres apart, and also thousands of tree species that are so rare that their niches cannot yet be quantified (Pitman et al. 2001, Honorio Coronado et al. 2009, ter Steege et al. 2020). Results on trees are often interpreted in such a way that dominant species have wide niches and due to this, they can be successful in a large variety of forest types (Pitman et al. 2001, 2013, Arellano et al. 2014). However, it is also true that adjacent Amazonian forest patches can have very different plant community compositions, and that distant patches with similar environmental conditions can have similar plant communities (Salo et al. 1986, Tuomisto et al. 1995, Honorio Coronado et al. 2009, Emilio et al. 2010, Fine et al. 2010). Broad-scale units with different soils and associated vegetation properties have also been recognised (Tuomisto et al. 2019). One possible way to reconcile these lines of evidence is that species that specialise to the most abundant kind of the available rainforest habitats become widespread, and appear generalists, even though they would not be able to reach high abundances in all forest types. Another possibility is that at least some of the broadly ranging species turn out to represent distinct, independently adapted lineages that may eventually turn out to be several geographically restricted species instead of a single broadly ranging one (Damasco et al. 2021).

This thesis rests strongly on niche theory. I focus on plant species responses to environmental variables, and on the role of the environmental variables in shaping forest structure and plant distributions. I use the community composition of plants in predicting soil variables. That said, I am aware that dispersal and biotic interactions also influence species distributions, and the roles of niche processes and dispersal processes are difficult to disentangle (Jabot et al. 2008). In simulations, both dispersal and environment drive changes in plant species relative abundances (Gravel et al. 2006), and in Amazonia they have been shown to do so in roughly equal proportions (Guèze et al. 2013). In this thesis, I use only species data and environmental data, without attempting to include dispersal or other biotic factors in the models. Also, analyses of phylogenies and speciation, often invoked in explaining niches, are beyond the scope of this thesis. I attempt to study the distribution of species relative to the environmental factors as it is seen currently, and I use the species as a static unit, even though in evolutionary timescales it is not static. However, in the indicator species practice, what is presently viewed as species can also be given indicator values, and I follow this practice here.

A few edaphically and hydrologically determined Amazonian forest types with characteristic physiognomy and plant community composition have been known for a long time. These include the swamp forests (Kalliola et al. 1991, Lähteenoja & Page 2011), floodplain forests (Assis et al. 2015), and forests on nutrient-poor white-sand soils (Anderson 1981, Adeney et al. 2016). These readily identifiable forest types cover a relatively small proportion of lowland Amazonia. The majority of Amazonian forests belong to the broad category of unflooded upland, or *terra firme*, forests on loamy and clayey soils. Broad-scale radar mapping projects have suggested differences in the physiognomy of *terra firme* forests as early as in the 1970s (RADAMBRASIL 1978), and traditional knowledge of local people also shows that *terra firm*e forests are not homogeneous (Shepard Jr. et al. 2001, Halme & Bodmer 2007, Abraão et al. 2008). During the last three decades, scientific research has increasingly begun to uncover substantial edaphic and floristic variation within the *terra firme* forests (Tuomisto et al. 1995, Phillips et al. 2003, Tuomisto et al. 2019). In this thesis, I concentrate on *terra firme* forests in the papers I–IV. Paper V is about pre-montane cloud forests, which do not belong to lowland Amazonian rainforests, but are inside the Amazon drainage basin.

Soil chemistry is one of the environmental factors that most often show congruent changes with plant species composition within the Amazonian *terra firme* forests (Phillips et al. 2003, Guèze et al. 2013, Tuomisto et al. 2016, Guevara Andino et al. 2021). Particularly, changes in the concentration of the base cations (the sum of Ca, K, Mg and Na) in soil has been shown to correlate with changes in plant community composition (Ruokolainen et al. 1997, Higgins et al. 2011, III, Zuquim et al. 2014, Cámara-Leret et al. 2017). This is likely because Ca, Mg and K belong to the essential plant nutrients, and their concentration varies considerably, up to two orders of magnitude, between different Amazonian *terra firme* sites (Sanchez & Buol 1974, Tuomisto et al. 2003a, Tuomisto et al. 2016, Hengl et al. 2017).

When considering the relative influence of different environmental variables on the plant community composition in Amazonia, one should take into account how the variables themselves vary at different spatial scales. Those factors that have high variation at the scale in question are likely to have the strongest influence. At small spatial scales, variation in the plant community composition is correlated with soil physical and chemical properties, such as grain size, nutrient concentration, pre-Columbian anthropogenic influence, hydrology and topography (Tuomisto & Ruokolainen 1994, Vormisto et al. 2004, Kinupp & Magnusson 2005, Zuquim et al. 2009, McMichael et al. 2014, Moulatlet et al. 2014, Marca-Zevallos et al. 2022). At large scales, there are broad trends in soil fertility that influence community composition (Tuomisto et al. 2016), but then also climate, particularly the amount and seasonality of rainfall, becomes important (ter Steege et al. 2006, Esquivel-Muelbert et al. 2016, Tuomisto et al. 2019, Guevara Andino et al. 2021).

1.2 Plants as indicators of soil properties

Since many plant species have niches that cover just a part of the soil gradient, the presence of a certain set of plant species (the species composition) at a site can be used as a proxy of soil properties (Diekmann 2003). Such bioindication is cost-effective because it is cheaper and less time-consuming than laboratory analyses of soil structure and chemistry.

In temperate and boreal regions, there is already a decades-long tradition of using indicator plants and plant species composition for soil inference (Cajander 1926, Kuusipalo 1985, Wilson et al. 2001, Gégout et al. 2003). It is particularly important in the field of forestry, where there is a need for estimating the capacity of timber production. Plants have also been used as bioindicators in e.g. mineral prospecting (Ahmad et al. 2022) and detection of archaeological sites (Pearson 1988). In Amazonia, the presence of certain fruit trees and palms can indicate pre-Columbian anthropogenic influence (Levis et al. 2012).

The development of a system of plant-soil bioindication for Amazonia is underway. The accumulation of botanical data, and the development of species distribution modelling during the past three decades have opened new research avenues to overcome the challenges posed by the low density of many species, and the large areas with sparse or nonexistent field data.

A significant part of the work on indicator species in Amazonia began in the 1990s by the University of Turku Amazon research team. Two particular groups of indicator plants have emerged: ferns (+ lycophytes) and Melastomataceae (Ruokolainen et al. 1997). Early studies showed that they can reproduce quite well the spatial pattern in the species composition of other plant groups, such as trees (Ruokolainen et al. 2007). The use of such surrogate groups greatly facilitates vegetation inventories in Amazonia. Tree inventories take a lot of working hours and specialist input due to practical difficulties in collecting tall plants, taxonomical problems and high species richness (Higgins & Ruokolainen 2004). Despite their obvious advantages, the efficacy of surrogate groups had only been tested on relatively small scales before the work published in paper **III**.

Melastomataceae is a large, predominantly tropical family of plants. Its species are shrubs, small trees and lianas. Various factors have led to their selection as a promising indicator group. They are relatively easy to collect, and to distinguish from other plant families by their characteristic leaf venation. Melastomataceae are consistently present in different kinds of Amazonian rainforests, and have a sufficiently large, yet manageable number of species, many of which have relatively narrow tolerances along the base cation gradient (Ruokolainen et al. 1997, **I, III**).

Ferns and lycophytes are ancient lineages of plants. Treelike ancestors from past environments are known, but presently most of the species are herbs. In the understorey of Amazonian rainforests, ferns are a prominent group (Costa 2004). They are promising bioindicators, since many relatively abundant fern species have narrow cation niches. The tiny spores of ferns are wind-dispersed, and due to their light weight, they can reach even remote locations with relative ease (Wolf et al. 2001). Because of this mobility and the fact that it is not likely that there would be large interspecific differences in the dispersal mode and/or capacity in ferns, the fern species composition at any Amazonian site can largely be expected to reflect the environmental conditions present at that site, and not dispersal limitation. Ferns as a group are also easily distinguished from other plants on the basis of their vegetative characteristics, which facilitates bioindicator surveys (Tuomisto 2006, Zuquim et al. 2014, **II**).

1.3 Modelling of plant species niches in Amazonia

The ecological niche is described by the factors that influence the performance of the individuals belonging to a species. Important environmental factors include e.g. water availability, nitrogen availability and annual average temperature. They can be seen to constitute a hypervolume (Hutchinson, 1957). The part of the hypervolume where a species is able to maintain positive population growth is its fundamental niche. Also biotic interactions such as competition, herbivory and mutualism shape the species' niche (Soberón 2007). They add to the multidimensionality of the niche and constrain it further. The result is the realised niche, which is expected to be smaller than the fundamental niche. In this thesis, I model the realised niches of Amazonian indicator species (Melastomataceae and *Adiantum* and *Lindsaea* ferns) along a single niche dimension: the concentration of base cations (Ca, K, Mg and Na) in the soil (**I**, **II**). I discuss the results in the context of community ecology and the practical usefulness of the species as indicators.

Informative niche properties include the species optimum, which is the point on the environmental gradient where the predicted abundance of the species reaches its maximum, and niche width, also called tolerance. The niche width can be variously defined, but it is meant to represent the part of the environmental gradient where the species is able to maintain a population. Some species are able to do that over a wide portion of the environmental gradient, whereas others are more specialised. Specialist species are said to have a narrow niche, whereas generalists have a wide niche. In this thesis, I use the terms “narrow niche” and “wide niche“ in the sense of amplitude along a single niche dimension. The choice of the niche width measure is arbitrary, and since I use occurrence data that produces the realised niche, the results do not necessarily imply that the species would actually be able to maintain positive population growth within the whole width of the estimated niche.

Knowledge about species niches can shed light on the mechanisms of community assembly, since differences in edaphic niches affect local species composition. Niche models can also help in the modelling of changes in species distributions under changing conditions. Species distribution models themselves are then useful in the planning of protected areas and sustainable use of species. Finally, the modelling of species responses to environmental gradients is important if one wishes to use the plants as indicators of environmental conditions.

Many Neotropical studies have reported that community-level patterns in plant species composition correlate with soil patterns (e.g. Duque et al. 2002, Phillips et al. 2003, John et al. 2007). There are less studies on the responses of individual species to soil variables. Preferences of species in relation to a classification of soil types or forest types has been documented for various plant groups (Pitman et al. 1999, Tuomisto & Poulsen 1996, Phillips et al. 2003, Salovaara et al. 2004, Fine et al. 2005, Cárdenas et al. 2007). Variation in species abundance along edaphic gradients has also been reported (e.g. Tuomisto & Ruokolainen 1994, Tuomisto et al. 2002). Actual modelling of edaphic niches with estimation of response curves and/or species optima and niche widths has been rarer, but it has been done at least for ferns (Tuomisto 2006, Zuquim et al. 2014), palms (Cámara-Leret et al. 2017) and trees (Toledo et al. 2012, Arellano et al. 2014).

The clear trends in the community level may be produced by a few abundant species that segregate along environmental gradients, while many rare species can simultaneously behave neutrally (Gravel et al. 2006). According to the niche breadth hypothesis (Brown 1984, Slatyer et al. 2013), geographically widespread species tend to have a wide niche. Local abundance has also been documented to correlate positively with both geographical range and niche width (McNaughton & Wolf 1970, Hanski 1982, Gaston & Lawton 1990).

Evidence for the role of niche size in the assembly of Amazonian plant communities is somewhat mixed. Some studies suggest that niche width and local abundance are positively correlated, and that geographically widespread species are also locally abundant (Pitman et al. 2001, 2013, Arellano et al. 2014), whereas others have challenged that view (Phillips et al. 2003, Tuomisto et al. 2003 a, Toledo et al. 2012). Most of these studies have been made with trees. Regional tree communities are often very species-rich, but dominated by a few common species, and most tree species are rare (Pitman et al. 2001, ter Steege et al. 2013, Draper et al. 2021). The set of dominant tree species differs between most Amazonian regions, i.e. dominant tree species of *terra firm*e rainforests are usually dominant only in one or two major Amazonian regions (ter Steege et al. 2013). Little is known about the preferences of rare trees, that constitute the majority of the species. Information on their niches is difficult to obtain due to their low densities: occurrence data accumulates very slowly with new vegetation surveys (Pitman et al. 1999). Modelling the niches of understorey plants, such as ferns and Melastomataceae, offers possibilities to compare niche characteristics of common and rare Amazonian species, since it is easier to obtain representative data on them than it is of trees. A tree plot of 1 ha typically contains 150–300 species and 450–500 individuals (≥ 10 cm diameter at breast height). A corresponding fern plot has on average 15–40 species represented by several thousand individuals. Due to the relatively high individuals per species ratio, reaching a sample size of relevant statistical power necessitates clearly less working hours with ferns than with trees. Moreover, the total number of species of trees is an order of magnitude larger than that of ferns. This means that species identification and taxonomical standardisation between different datasets needs much more specialist input when working with trees (e.g. Honorio Coronado et al. 2009).

Niche modelling can also provide insights into how commonly the species responses deviate from the theoretical Gaussian response, i.e. are skewed or non-unimodal (Austin 1976, Austin & Smith 1989, Lawesson & Oksanen 2002, Oksanen & Minchin 2002, Rydgren et al. 2003). Toledo et al. (2012) found unimodal responses to be rare among Amazonian rainforest trees, and they termed the result surprising since the environmental gradients covered were relatively long. On an infinitely long gradient, all responses are theoretically unimodal. The large datasets I use in papers **I** and **II** comprise a long soil base cation gradient, and therefore offer a possibility to test this with other Amazonian plants.

In Amazonia, there are indications that understorey plants tend to have narrower soil niches and higher species turnover between habitats than canopy trees do (Ruokolainen & Vormisto 2000, Duque et al. 2002), possibly due to their more limited dispersal abilities (Kristiansen et al. 2012, Arellano et al. 2014), or the fact that high-statured plants have a larger root sphere and gather nutrients from a larger area. In several taxonomic groups, high-statured tropical rainforest plants appear to be ecologically more generalist than smaller plants (Webb 1967, Duque et al. 2002, Chust et al. 2006), but these indications might also be due to the fact that it is more difficult to obtain spatially representative samples of large canopy trees than of understory plants (Jones et al. 2008). Some studies also show an inverse relationship in the degree of environmental control: the distribution of tall species is more correlated with environmental variation than that of understorey species (Paoli et al. 2006, Guèze et al. 2013). Nevertheless, among Amazonian palms it has been shown that large canopy species tend to have wider niches than small-statured understory species (Ruokolainen & Vormisto 2000, Kristiansen et al. 2012, Cámara-Leret et al. 2017).

When modelling plant species niches at the scale of the whole Amazonia, one must take into account that the niche width of geographically widespread species can differ between regions. This can happen for several reasons. The niche can become narrower towards the margins of the range because other environmental constraints restrict species performance there and the species can only persist on the most favourable part of the gradient. The niche can also expand towards the range margins due to competitive release, if the competitors are not able to maintain populations there. The niche can shift to a different position on the gradient due to local adaptation. For example, if a larger proportion of the sites available to the species are low in nutrients, natural selection may lead to a niche shift towards the low-nutrient part of the gradient in that region (Hájková et al. 2008). In Europe, regional variation in the soil niches of plants takes place along the south-north climatic gradient (Reinecke et al. 2016, Hedwall et al. 2019). Climatic variation exists also in Amazonia. There are regional differences particularly in the rainfall seasonality (de Moura et al. 2015). In addition to local adaptation, speciation can be underway, and in some cases, a widespread generalist species can turn out to be several geographically restricted and perhaps locally specialised species (Damasco et al. 2021). Knowledge on possible geographical differences in species niches can improve the modelling of niches and species distributions in Amazonia.

1.4 Soils and forest structure in the Amazonian lowland rainforests

Many tropical rainforests have a complex, multilayered structure. Despite the highly variable edaphic properties and species composition (Tuomisto et al. 1995, Tuomisto et al. 2003a), most Amazonian rainforests look superficially similar – a common metaphor is a uniform field of broccoli. The differences in edaphic properties and in species composition between different areas and between forest types appear much larger than the corresponding differences in forest structure.

Soils have direct influences on plant species composition via the physiological responses of species, but soils can also influence species composition indirectly through forest structure. Canopy density can differ between forests that grow on different soils. A denser canopy means a shadier understorey environment, which is important for both understory herbs and shrubs as well as for young individuals of canopy trees (Nicotra et al. 1999, Montgomery & Chazdon 2002). The structure of the canopy is important for the rich community of epiphytic plants (Ozanne et al. 2003). Stratification increases the availability of microenvironments, contributing to the diversity of many kinds of organisms (MacArthur & MacArthur 1961, Erwin 1982, Scheffers et al. 2013). Understorey structure is likely to be an important variable for birds (Pomara et al. 2012). Forest structure in general is thus also an important driver of evolution, community composition, and of regional differences in species' functional traits. Quantifying the structural differences could help explain differences in these properties between forest types. In papers **IV** and **V**, I attempt to quantify differences in forest structure and study its linkage to species composition.

Differences between forests growing on different soils are discernible even on satellite imagery (Higgins et al. 2011, 2012). The soil itself, from below the layers of canopy and understorey plants, can hardly have any direct influence on the reflected light captured by a satellite's sensors, but patterns of canopy reflectance are still congruent with edaphic patterns in Amazonia (Higgins et al. 2012, Sirén et al. 2013). This has mostly been attributed to differences in species composition, which causes variation in the chemical or physical characteristics of the leaves, but it can also be due to differences in the structure of the canopy (Lu et al. 2004).

In Brazilian forest maps, two main types of non-inundated *terra firme* rainforests are recognized: floresta aberta (open forest) and floresta densa (dense forest; [Veloso 1974, IBGE 2004]). They are named for the perceived differences in their structure, but the differences had not been quantified at the time of the field work for paper **IV**, in 2010. In Peru, the forests growing on the Pebas and Nauta geological formations appear to correspond to the open and dense forests in Brazilian forest maps, respectively (Higgins et al. 2011). It is possible to identify them via manual interpretation of enhanced satellite images (Higgins et al. 2012). The two formations differ in the concentration of nutrient cations and in soil grain size. The clayey Pebas Formation is on average approximately ten times richer in cations than the loamy to sandy Nauta Formation, although individual locations on both formations can have similar cation concentrations (Hoorn 1993, Rebata et al. 2006, Higgins et al. 2012). They are also floristically distinct (Ruokolainen & Tuomisto 1998, Fine et al. 2005, Higgins et al. 2011).

When planning the field sampling for paper **IV**, it was possible to hypothesise that on the more nutrient-rich Pebas formation, the forests could be more open than the forests on Nauta formation, since forests on more fertile soils in Amazonia were known to have a faster tree turnover rate than forests on poorer soils (Phillips et al. 2004, Quesada et al. 2012). A faster turnover rate would lead to more frequent gap formation. On the other hand, a higher growth rate of saplings on nutrient-rich soil could lead to faster closure of gaps, so after all, it was not completely clear what kind of difference in understorey structure to expect.

1.5 Forest structure, microclimate and epiphytes in the Andean pre-montane cloud forests

The structure of the pre-montane cloud forests of the eastern slopes of the tropical Andes differs from that of lowland Amazonian forests. Cloud forests have a lower canopy, less emergent trees, and a higher epiphyte load (Bruijnzeel et al. 2011, **V**). Epiphytic communities are intimately linked to forest structure, since the epiphyte community is entirely dependent on trees.

In the context of Amazonia, the pre-montane cloud forests are particularly vulnerable to climate change since they occupy a narrow altitudinal band between the extensive Amazonian lowlands and the treeless páramos of the high Andes. The upward shift of biotic altitudinal zones is threatening particularly the uppermost cloud forests as they eventually will have nowhere to migrate. Deforestation and global warming often have similar consequences to cloud forests: the conditions become drier (Zotz & Bader 2009). Deforestation due to human activities has been more rapid in the pre-montane cloud forests than in most of the lowland Amazonian forests, due to the higher human population density in the Andes compared to that in lowland Amazonia. According to some estimates, around 90 % of original cloud forests have already been deforested (Gradstein 2008), a staggering percentage compared to the 17 % of lowland Amazonian rainforests (Lovejoy & Nobre 2019).

Pre-montane cloud forests are important producers of clean water. Their canopy intercepts rain and slows it down on its way to the ground. In this way, these forests protect the steep terrain from erosion by decreasing throughfall and subsequent surface runoff (Veneklaas & van Ek 1990). They also increase the infiltration of rainwater into groundwater. In the absence of these forests, torrential rains often lead to landslides on steep slopes. Erosion and sedimentation increase as deforestation proceeds, leading to growing turbidity of river waters, more pronounced fluctuations in river discharge, and a dropping groundwater level. Against this background, research on the mechanisms that affect the hydrological properties of cloud forests is urgently needed. Epiphytic plants have an important role in the hydrological cycle because they intercept and retain horizontal precipitation (water from fog and clouds). They also capture a significant portion of rainfall, which they subsequently transpire back into the atmosphere, contributing to the maintenance of the precipitation regime (León Vargas et al. 2006, Gradstein 2008, Tobón et al. 2010).

Epiphytes are plants that germinate and complete their life cycle on the trunks and branches of trees. They are most abundant on the canopy branches, but they do occur in all strata of the forest, including on the tree trunks, just above ground level. Epiphytes collect water and nutrients from rainfall, falling debris, and water flowing along the trunks and branches of trees. They are not parasites; they do not extend their roots to the tissues of the host tree but use the tree only for mechanical support. In tropical rainforests, epiphytes are species-rich and abundant, and they belong to many plant lineages, from mosses and ferns to orchids and bromeliads (Zotz 2013). They are particularly species-rich in cloud forests, where cloud condensation provides high levels of atmospheric humidity as the warm and moist air from the surrounding lowlands ascends the mountains, creating favourable conditions for epiphytic way of life (Zotz & Bader 2009).

Forest structural characteristics have direct influences on epiphytes. Important structural characteristics include canopy openness, tree basal area and average tree height.

A closed canopy provides shade and is beneficial for those epiphytic species that are specialised to the lower trunks and branches. Closed canopy also contributes to the maintenance of high within-stand humidity. On the other hand, a more open canopy allows higher light penetration to the lower forest strata, which can be beneficial for some light-demanding epiphyte species. A more open canopy also leads to higher within-stand temperatures. If predictive models of epiphyte abundance are based on forest structural variables only, the inclusion of these microclimatic variables should improve the models. This is because forest structure has an indirect influence on the epiphytes via microclimate, but the influence of microclimate on epiphytes is direct.

Basal area and average tree height can generally be expected to be positively related to epiphyte abundance, since the available habitat area for epiphytes increases with them. Larger trees have more surface area for epiphyte attachment. Larger trees also tend to be older than small trees, which means that the epiphytes have had more time to colonise the trunks and branches. Epiphyte species richness tends to be higher in mature cloud forests than in secondary forests, due to their more complex structure that leads to a more suitable microclimate and a larger number of possible epiphyte attachment sites (Nöske et al. 2008).

Deforestation, fragmentation and selective logging lead to higher temperatures within the remaining forest stands due to the increase in the percentage of open areas that receive direct sunlight all the way to the ground level (Krömer & Gradstein 2004). More intensive heat causes loss of the high humidity, pivotal to the rich epiphyte flora of the cloud forests, leading to a decrease in epiphyte cover and epiphyte species richness (Nadkarni & Solano 2002). Selective logging often targets large trees, reducing available epiphyte habitat (Zotz & Bader 2009). Anthropogenic influence tends to reduce basal area and average tree height, and increase canopy openness.

Epiphytes can be divided into two functional groups: vascular and non-vascular epiphytes. Vascular epiphytes belong to various vascular plant groups, of which the aroids, bromeliads, ferns and orchids comprise the great majority of all species (Gentry & Dodson 1987, Gradstein 2008). Filmy ferns, bryophytes and lichens are non-vascular epiphytes. Of these three non-vascular groups, this study only considers bryophytes, which can be the largest component of epiphytic live biomass in cloud forests (Nadkarni 1984, Gehrig-Downie et al. 2011).

Up in the canopy, water is often a limiting resource. Vascular epiphytes have many anatomical and physiological adaptations that enhance their drought tolerance. These range from the succulent and sclerophyllous leaves of aroids, to the water storage bulbs of orchids and the water-absorbing trichomes and water-collecting leaf tanks of bromeliads (Mantovani 1999). The waxy cuticle of vascular epiphytes protects them from desiccation, and the stomata of their leaves can regulate transpiration. Many bromeliads do CAM-photosynthesis with nocturnal CO_2 absorption, keeping their stomata closed during the intense daytime sunlight (Males 2016).

Bryophytes, to the contrary, lack the regulatory adaptations of vascular epiphytes, and desiccate easily. They respond within minutes to the wetness or dryness of their surroundings, and their photosynthesis ceases when the relative air humidity falls below approximately 95% (León Vargas et al. 2006). However, they are able to survive prolonged dry periods in an inactive state and can return to positive carbon balance within an hour or two after getting hydrated again (León Vargas et al. 2006, Zotz & Bader 2009). High temperatures also limit the performance of bryophytes. High daytime temperatures can lead to the relative air humidity dropping below the critical threshold of 95%. High night-time temperatures can increase respiration while the bryophytes are not photosynthesizing, to the extent that their net carbon balance becomes negative.

These differences in traits between vascular epiphytes and bryophytes have dual consequences. First, bryophytes are likely to be more sensitive to the within-stand air humidity intermittently dropping below 95 % during daytime, making their cover percentage a reasonable proxy for humidity conditions within pre-montane cloud forests (Karger et al. 2012). They are also likely to be more sensitive to high temperatures than vascular epiphytes are. Second, vascular epiphytes might be more affected by occasional, prolonged dry periods than bryophytes. Most adult vascular epiphytes may survive even the longest dry periods well, but young individuals will perish if the drought lasts for long enough; most vascular plants cannot recover from desiccation in the same way as bryophytes do.

It is thus likely that vascular and non-vascular epiphytes respond in different ways to variation in forest structure and microclimate. The different groups of vascular epiphytes are also likely to have variable responses due to their different functional traits. Ferns, in particular, differ from seed plants in having less optimised leaves for environments where light intensity and evaporation are high (Zwieniecki & Boyce 2014).

The way how forest structure and microclimate interact in regulating the abundance of different epiphyte groups in tropical pre-montane cloud forests is not completely clear. Forest structural variables are inexpensive and straightforward to measure. Using them as proxies for microclimate could facilitate surveys. A better understanding of the role of forest structure and microclimate as drivers of epiphyte abundance could help in protecting these vulnerable environments, and in predicting their future responses to anthropogenic disturbance and climate change.

1.6 Aims of the thesis

The general aims of this thesis are to increase knowledge about the relationship that Amazonian plants have with soils, and to explore possibilities to utilise this relationship in bioindication. I also study the influence that soil has on forest structure, and the influence that forest structure has on epiphytic plants.

I have worked on this thesis part-time and intermittently for more than a decade. The aims of the thesis, and the possibilities of working on it, have changed along the way, and the resultant combination of original papers reflects this history.

Bioindication can be improved if the precise form of the plant species' relationship to soil properties can be modelled. Until recently, large enough datasets for this have hardly been available in Amazonia. In papers **I** and **II**, I make use of large Amazonian-wide datasets of Melastomataceae, and of two widespread and common fern genera, *Adiantum* and *Lindsaea*. I characterise their realised soil cation niches, specifically their response shapes, optima and niche widths. I also assess factors that correlate with the niche properties of common and rare species.

In paper **III**, I test the possibilities of bioindication of various soil properties using Melastomataceae. I also compare two different techniques for prediction: k-NN and weighted averaging.

In paper **IV**, I make an attempt to quantify differences in understorey structure between Amazonian *terra firme* rainforests growing on two widespread geological formations: the Pebas Formation and the Nauta Formation. I also investigate the local distribution of a frequent and abundant palm species, *Iriartea deltoidea*, on these geological formations.

In paper **V**, I investigate how forest structure affects the abundance of vascular and non-vascular epiphytes in the tropical pre-montane cloud forests of northern Peru. I also attempt to clarify how forest structure and microclimate interact, and how different groups of epiphytes respond to their variation.

2 Methods

2.1 Sampling strategies and data collection

The large datasets in papers **I–III** have been collected during three decades, by research groups based in Brazil and in Finland. The field sampling sites are spread over large areas of Amazonia (>1000 km between most distant sites). In papers **IV** and **V**, the focus is more localised (<100 km between most distant sites), and I was personally in charge of data collecting (Figure 1). I collected the data for paper **IV** with a small team of local helpers, and the data for paper **V** in collaboration with biologists Johanna Toivonen and Gabriel Trujillo (Table 1).

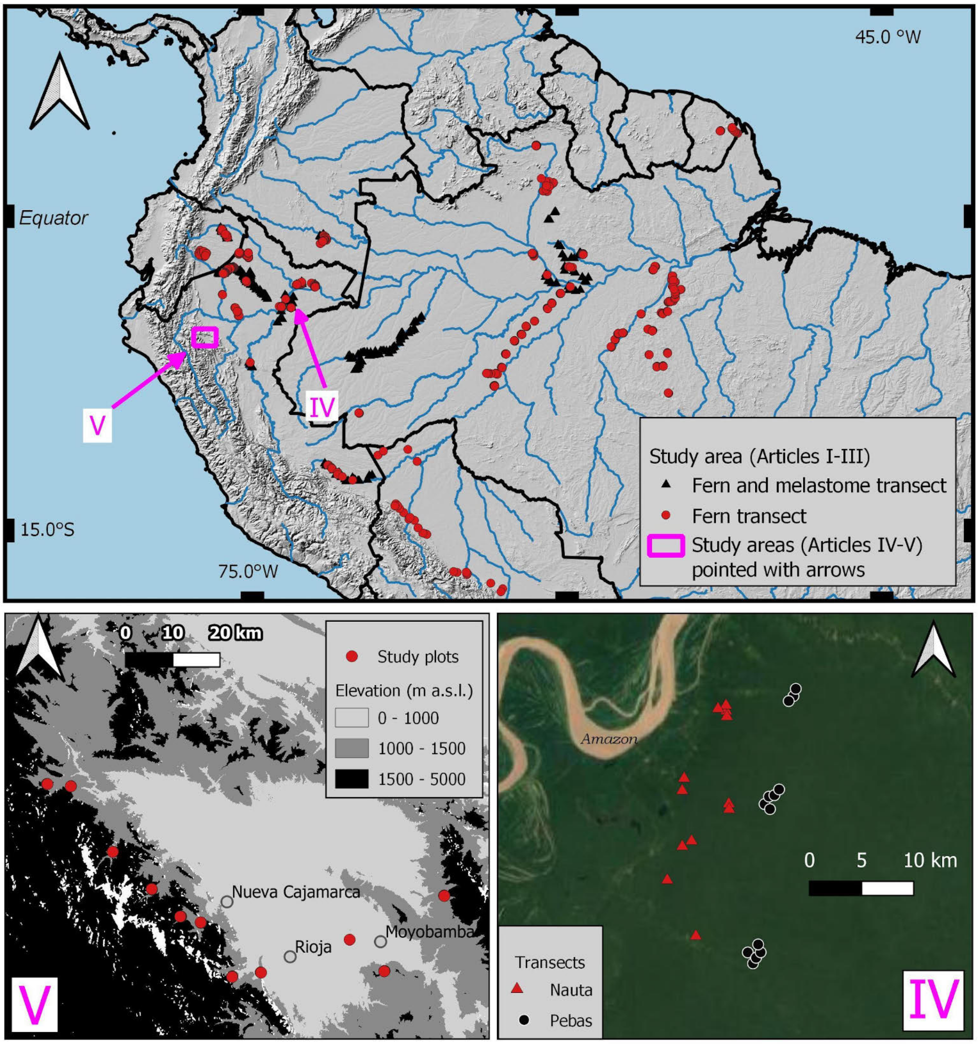

Figure 1. Map of the study areas. More detailed maps of each study area are in the original papers.

2.2 Species data

The species data for papers **I–III** has been collected using two different sampling strategies, performed by two different research groups: the PPBio group (Brazilian program for biodiversity research, explained in more detail in e.g. Zuquim et al. 2012) and the UTU-ART group (University of Turku Amazon research team, explained in more detail in papers **I** and **II**). Methodological differences in the data collection are minor, the largest ones being the dimensions and topographical orientation of the transects. For Melastomataceae, only data collected with the UTU-

ART methodology is available. For *Adiantum* and *Lindsaea* ferns (paper **II**), I use both kinds of data jointly.

In the UTU-ART methodology, an inventory transect with dimensions 5 m x 500 m is established following a predetermined compass bearing. In the PPBio methodology, a transect with dimensions 2 m x 250 m is opened following the terrain contour line (staying at the same elevation throughout the transect).

In the papers **I** and **III**, I use the Melastomataceae data in its entirety (all species and morphospecies included). In paper **II**, I use the *Adiantum* and *Lindsaea* data, but restrict the analysis to named species only, not including any morphospecies.

All Melastomataceae individuals were identified to a named species (43 %) or unnamed morphospecies (57 %) by Kalle Ruokolainen (in continuation I use the term ‘species’ for both categories). All *Adiantum* and *Lindsaea* individuals were determined to a named species, and the taxonomy was standardised throughout the UTU and PPBio datasets by Hanna Tuomisto.

There are 1061 transects in total. The majority of them have presence-absence records. Species abundance was recorded in a subset of 402 transects (Table 1).

Table 1. Characteristics of the data used in each paper.

ORIGINAL PUBLICATION OR MANUSCRIPT	SAMPLING DESIGN	COUNTRY	DATA TYPE	FIELD DATA COLLECTORS (INITIALS) *
I	284 melastome transects (5x500 m) in northwestern, southwestern and central Amazonian lowlands	Brazil, Colombia, Ecuador, Peru	Melastome abundance, soil samples (cations)	KR, HT
II	1061 fern transects (5x500 m or 2x250 m) in northwestern, southwestern and central Amazonian lowlands	Bolivia, Brazil, Colombia, Ecuador, French Guiana, Peru	Fern and lycophyte presence, presence-absence and abundance, soil samples (cations)	HT, GZ, GC, GM, PW, AA, AS, EP, SL, MH
III	311 melastome transects (5x500 m) in western Amazonian lowlands	Colombia, Ecuador, Peru	Melastome presence-absence, soil samples (Al, cations, pH, P and loss on ignition)	KR, HT, NL, MH
IV	25 transects (500 m long) in one locality in lowland western Amazonia	Peru	Forest structure, most abundant ferns, *Iriartea* palm, satellite imagery	LS
V	36 square plots (20 * 20 m) in 11 sites in the pre-montane cloud forests	Peru	Forest structure, epiphyte abundance, microclimate	JT, LS, GT

* AA= Alfonso Alonso, AS= Anders Sirén, EP= Eneas Pérez, GC= Glenda Cárdenas, GM= Gabriel Moulatlet, GT= Gabriel Trujillo, GZ= Gabriela Zuquim, HT= Hanna Tuomisto, JT= Johanna Toivonen, KR= Kalle Ruokolainen, LS= Lassi Suominen, MH= Mark Higgins, NL= Nelly Llerena, PW= Patrick Weigelt, SL= Samuli Lehtonen.

2.3 Forest structure measurements

I conducted forest structure measurements in the lowland rainforest of northern Peru (**IV**). I decided to measure understorey structure because that can be accomplished with simple equipment: a compass, a hand-held GPS, a measuring tape and a canopy-scope plate are basically all the research tools needed in the field. I reasoned that if marked differences in understorey structure would be detected in my simple and inexpensive way, it would be a strong indication of these two floristically different forest types also being structurally different, and could help in designing and targeting future investigations. Also, exact measurements of understorey stem density can be obtained only with ground-based measurements, not with remote sensing. I also took into account the smallest stems (<2,5 cm DBH [diameter at breast height]), which are ignored in many studies.

I established inventory transects of 500 m, 12 of them on Nauta formation and 13 on Pebas formation. The transects were similar to the fern and Melastomataceae transects of the UTU-ART group: each transect was opened following a predetermined compass bearing, trespassing valleys, hills and treefall gaps alike. I measured forest structure at observation points (which were 10 m off the transect line, on alternating sides of the transect) in the beginning of the transect and then at 50 m intervals until the end of the transect. This yielded a total of 275 observation points.

At each observation point, I estimated the stem density of three different size classes of trees: saplings (< 2.5 cm DBH), poles (≥ 2.5 ≤ 10 cm DBH), and trees (> 10 cm DBH). I estimated canopy openness with a canopy-scoping method (Brown et al. 2000). I also collected and photographed three most abundant fern and lycophyte species per observation point. This was done to verify that the observation point actually was located on the intended geological formation, which had been determined beforehand based on a geological map and satellite imagery. I also surveyed the presence of the palm *Iriartea deltoidea* Ruiz & Pav. on the two different geological formations.

The sampling design of the study on forest structure and epiphytes in the pre-montane cloud forests (paper **V**) differs from the rest of the papers in this thesis. Instead of line transects, which are impractical in the often precipitously hilly pre-montane forests, I used rectangular (20 x 20 m) plots. To cover a sufficient area per site, I established rows of four plots (in total 0.16 ha; on some sites the plots were fewer than this due to difficult terrain). In total I established 36 plots.

I measured the following variables at each plot: Forest structure, bryophyte cover, and abundance of vascular epiphyte groups (aroids, bromeliads, ferns, orchids and palms). The forest structural variables recorded were the height and circumference of all trees > 10 cm DBH (diameter at breast height), canopy openness (an average of five canopy-scope measurements per plot). The former measurements allow calculating tree basal area (in square metres) and average tree height per plot, whereas the latter provides a proxy for the intensity of light in the forest understorey. Bryophyte cover was estimated both visually at the level of the whole plot and systematically with a sampling grid from the lower trunks of a subset of the trees. Abundance of vascular epiphytes was estimated visually by counting individuals on the tree trunks from the ground level up into the level of lowermost canopy branches, using binoculars if needed.

2.4 Environmental data

Soil samples were collected in all fern and Melastomataceae transects. Soil samples were taken in several locations along each transect, and they were subsequently pooled to represent the average soil conditions in the transect. Several soil properties were analysed from the samples, most importantly the concentrations of the base cations. In papers **I** and **II**, I work only with the sum of base cation concentrations (Ca, Mg, K, Na). In paper **III**, I use Melastomataceae species composition to predict several soil variables.

In paper **IV**, I used a Landsat image to estimate whether the reflectance values of the forests on Nauta and Pebas formations were correlated either with the forest structural variables or with floristic composition (in this case, represented by ferns and lycophytes). I also used the Landsat image to evaluate how representative the 25 field transects were of the Nauta and Pebas formations in general.

In paper **V**, microclimate was recorded in a subset of 17 plots. We placed automatic sensors within the plots to record temperature and humidity data every 1 hour between January and May 2015. This period represents the core rainy season and the transition to the dry season. From the sensor readings we extracted values describing minimum humidity and maximum temperature. We considered these extreme values to be more important to epiphytes than average values.

2.5 Data analysis

2.5.1 Species response modelling and the prediction of soil properties

Species niches can be studied based on observational data in various ways, but a common denominator for all these approaches is that they attempt to uncover the realised niche (Soberón 2007, Devictor et al. 2010). I chose the approach of modelling species responses on a single niche dimension. The advantage of the approach is that the results tell explicitly about both niche breadth and niche position along the selected dimension (in my case, the concentration of base cations in the soil). Disadvantages are e.g. the fact that true niches are multidimensional, and other environmental variables probably affect the species response along the selected dimension in unknown ways (Soberón 2007).

A species' response to an environmental variable is commonly visualised by plotting the environmental variable on the horizontal axis of a two-dimensional graph, and the abundance of the species on the vertical axis. The abundance observations (individuals per unit area) form a cloud of points. In this cloud, a curve that minimises the distance to the points can be fitted in various ways. The result is a species response curve, classically bell-shaped (Gaussian), but in some cases of other kinds as well (Whittaker 1960, Brown 1984, Austin et al. 1984, Bruno et al. 2003). The response curve shows how the abundance of the species is predicted to vary along the selected niche dimension. Various niche characteristics, such as the species optimum and niche width, can be derived from the response curve.

There is a wealth of methods of estimating a plant species' response to an environmental gradient. I chose to use HOF models and weighted averaging.

The HOF models (Huisman et al. 1993, Lawesson & Oksanen 2002, Oksanen & Minchin 2002, Jansen & Oksanen 2013, **I**, **II**) offer a way of parametrically choosing the most suitable model from a predefined set of models. Since the criteria for model selection are defined mathematically, the degree of arbitrariness in model choice is reduced. The other side of the coin is that the model shapes are limited to the predefined set. This set, however, is quite comprehensive. It includes seven possible model shapes. In the order of increasing complexity, the modelled response can be: no response (flat), monotonically increasing or decreasing, plateau (response increases towards one end of the gradient, but a plateau is reached before the end), unimodal symmetric, unimodal asymmetric, bimodal symmetric or bimodal asymmetric. The actual occurrences of each species along the gradient are fitted to these models. Parsimony is used in model choice, i.e., if two models fit the data equally well, the simpler of them is chosen.

HOF models can be calculated based on data on species presences and absences. Data on species abundance can help in producing more accurate models, but on the other hand, models based on abundance data are more sensitive to large intraspecific differences in abundance between sites, a frequent situation in plant ecology. In paper **I**, I use only abundance data for estimating Melastomataceae species responses, whereas in paper **II**, I test HOF-modelling with both presence-absence and abundance data of *Adiantum* and *Lindsaea* ferns. The use of presence-absence data is tempting because it is faster to collect than abundance data and is also more widely available.

In weighted averaging (ter Braak & van Dam 1989, Birks et al. 1990), the species optimum is defined as the weighted average of the environmental value in all those sites in which the species occurs, with the species abundances used as weights. Niche width, then, is defined as one standard deviation above and below the optimum. No actual curve is drawn, but a symmetric unimodal (bell-shaped, Gaussian) response is assumed. In paper **II**, I use weighted averaging with presence-only, presence-absence, and abundance data on *Adiantum* and *Lindsaea* ferns.

I compared species niche positions estimated with HOF modelling and weighted averaging between different regions, to find out how consistent the niche estimates are. I did this by calculating Pearson correlations between species optima, and average overlap between species tolerances (using adaptations of the Sørensen index and Simpson's index) based on the regional datasets. In the case of ferns, I used three different regional sets of transects: northwestern, southwestern and central Amazonia. Soil-wise the two western regions share an Andean affinity, whereas in central Amazonia, the soils have their origin in the shield regions of Mato Grosso and Guyana. The northwestern region has a nearly aseasonal climate, whereas rainfall in the two other regions is seasonal. The Melastomataceae data has a smaller regional extent than the fern data. I compared Melastomataceae niches between two regional datasets: northwestern (NW) Amazonia, and southwestern (SW) and central Amazonia combined.

Various aspects of species commonness can covary with niche width. These aspects include species frequency, total abundance, local abundance, and geographical range. I estimated them as follows: Frequency is the number of transects in which the species occurs; total abundance is the number of individuals in the abundance-quality transects; local abundance is the number of individuals per transect containing the species (density when present); and geographical range is the largest geographical distance between any two transects containing the species. I calculated the R^2 for a model between all these aspects of commonness and niche width for both plant groups, except for geographical range, which I did not estimate for Melastomataceae due to the more restricted spatial extent of their sampling.

Since many plant species only occur on a certain kind of soil, the set of plant species found on a site can be used to predict the values of soil variables at that site (Kuusipalo 1985, Wilson et al. 2001, Gégout et al. 2003). In paper **III**, I use two simple techniques for prediction: *k*-nearest neighbour estimation (*k*-NN) and weighted averaging calibration, and compare their accuracy.

In *k*-NN, the *k* stands for the number of neighbours. In this study, the neighbours to any focal site (i.e., the site for which the predictions are made) are those other sites where the species composition is most similar to the focal site. Thus, being a neighbour is about floristic proximity and not necessarily geographical proximity. The idea is that for the *k* neighbouring sites, also soil properties have been laboratory analysed and are thus known, whereas for the focal site, they are not. If only a single neighbour is used ($k = 1$), then the soil variable value in that neighbour is used as the prediction for the focal site. If more neighbours are selected, the weighted average of the soil variable among the *k* neighbours is the prediction, with the nearness criterion used as the weight, i.e. floristically more similar sites have more weight than less similar sites when calculating the predicted value. As *k* increases, the prediction converges towards the mean of that soil variable in the entire dataset. I used a database of ca. 300 Melastomataceae and soil transects in such a way that each transect was used as the focal site in turn. Since the measured soil values for each site were known, I was able to compare the measured soil values with the predicted soil values by calculating R^2, root-mean-square error (RMSE) and maximum bias, and thus estimate the accuracy of the *k*-NN method (Katila & Tomppo 2001, **III**).

Calibration via weighted averaging (ter Braak 1987, **III**) differs from *k*-NN in such a way that it is based on the species' indicator value. A species' indicator value for any soil variable is calculated as the weighted average of the soil variable values at all sites where the species occurs. The weighing means that species abundances are used as weights when calculating the average. In paper **III**, I used only presence-absence data, so in practice, the indicator value was simply the arithmetic average of the soil value at all sites containing the species. The predicted soil values for any focal site were arithmetic averages of the indicator values of all species present on the focal site. I calculated indicator values separately before calculating the predicted value for each focal site, so that the soil values in the focal site itself did not enter the calculation of indicator values. This is also called leave-one-out cross-validation (ter Braak & Juggins 1993).

2.5.2 Delineation of rarity categories

In paper **II**, I estimate species rarity and assess its relationship with niche properties. I apply Rabinowitz's (1981) rarity categories. These categories form a grid of eight cells (Table 3 in the Results) that classifies species based on three variables: geographic range (large or restricted), niche specificity (narrow or wide niche), and local abundance (locally abundant or locally scarce). Seven out of eight grid cells represent some form of rarity. One cell holds the species that are not rare: the geographically widespread, locally abundant generalists.

These rarity categories have been used for plants in different geographical areas (Caiafa & Martins 2010, Choe et al. 2019), also for trees in southwestern Amazonia (Pitman et al. 1999). I apply the same framework to ferns with spatially extensive sampling, including abundance data, which opens a unique possibility to study commonness and rarity in Amazonian plants. I assess the relative proportions of different kinds of rarity, their correlation with niche properties, and compare the results with earlier studies on trees. Rabinowitz's rarity framework can also be used in the search for indicator species. The best indicators belong to the category defined by a large range, high local abundance, and a narrow niche. In other words, a good indicator species is widely distributed, easy to find in suitable sites, but restricted to a certain kind of habitat.

Using rarity categories involves a few arbitrary decisions about the borders between the categories (i.e. choosing the limits between large and small geographical range, narrow and wide niche, high and low local abundance; Pitman et al. 1999). If the variables are measured on a continuous numerical scale, averages can be used as class limits (Choe et al. 2019). If the variables are categorical, an informed choice can be made based on what is known about the local species distributions (Caiafa & Martins 2010). An obvious drawback of using averages or medians in delineating Rabinowitz categories is that all categories are then deemed to have some species. I ended up using the median as the limit criterion between high and low local abundance. I reasoned that the median is a better indicator than the average, due to the vast interspecific variation in abundance: a small proportion of the species were locally very abundant, which inflates the average. Using average would have made the "locally abundant" category too restrictive. For niche width, I used 25% of gradient length as the limit between narrow- and wide-niche species, a limit that I decided based on the set of HOF models made for all species. Species that have a niche width that spans a quarter of the gradient or less are promising indicator species. I used 1000 km between the two most distant observations as the limit between a large and a restricted geographical range. This is because a species with a range smaller than 1000 km between the two most distant observations is reasonably likely to be confined to only one of the Amazonian regions (northwestern, southwestern, or central Amazonia). This limit for a small range is an order of magnitude larger compared to the limit for tropical plant range endemism proposed by Gentry (1986). Selecting a smaller limit would, however, have left the narrow range category with very few species. This probably reflects the relative lack of dispersal limitation in ferns.

2.5.3 Quantification of forest structural differences

In paper **IV**, I compare the distribution of forest structural variables in the Nauta and Pebas formations with parametric *t*-tests and nonparametric permutation tests. Both tests can be used to address the same basic question: how likely is it to end up with a similar difference in means by chance? Since the sample size was relatively small, I also did a test of statistical power for the *t*-tests to answer the interesting question of how much larger the sample size should have been in order to detect a statistically significant difference in those cases where the difference in means was nonsignificant. The statistical power, or likelihood of detecting a significant difference in means, is dependent not only on sample size but also on how large the difference in means is and how much variation there is around the means. In the power tests, I assumed all other variables except sample size to be constant.

Since several previous studies had shown that floristic composition is often correlated with reflectance values in Landsat image pixels (Tuomisto et al. 2003 b, Salovaara et al. 2005, Higgins et al. 2012), I wanted to know whether this would also be the case with forest structural variables. I also wanted to amplify the spatial coverage of the study with remotely sensed information. I used multiple regression to construct a model of the forest structural variables that correlate significantly with the reflectance values within windows centred on the field transects. I applied this regression model to predict forest structural variables in >2000 reference windows of similar size, placed elsewhere on the same forest types, on the same satellite image. In that way, I was able to compare much larger surfaces of the two formations than with the field sampling alone. I also used the Landsat image to estimate how representative the field sampling was of the Nauta and Pebas formations more broadly. To this end, I calculated six-dimensional spectral spaces based on both the reference windows and windows centred on the field transects. I used the Euclidean distances between the spectral space centroids of the reference transects and field transects to estimate whether using my field transects would lead to under- or overestimating the regional difference between the formations.

2.5.4 The role of forest structure and microclimate as drivers of epiphyte abundance

In paper **V**, my aim was to clarify the roles of forest structure and microclimate in controlling the abundance of epiphytes in the pre-montane cloud forests. I assessed the relationship between epiphyte abundance, microclimate and forest structural variables with regression analysis, where epiphyte abundance was the dependent variable and forest structure and microclimate were the explanatory variables. To be able to use each individual plot as an independent observation even if the plots were within the same site, I carried out simultaneous autoregressive (SAR) models that account for spatial autocorrelation (Beale et al. 2010, Kissling & Carl 2008).

Due to the limited number of data points, it was not possible to construct a regression model with more than two explanatory variables. Therefore, I tested all possible pairs of explanatory variables from the set of five variables (minimum humidity, maximum temperature, canopy openness, pooled basal area, average tree height), and then used multimodel inference to rank the variables according to their importance (Symonds & Moussalli 2011). I used the AICc statistic (Akaike Information Criterion corrected for small sample sizes) to choose the best approximating model among the candidate models.

3 Results and Discussion

3.1 Species optima and response shapes on the base cation gradient

All studied *Adiantum* and *Lindsaea* ferns (**II**) and 99.5 % of Melastomataceae (**I**) showed statistically some kind of soil base cation preference. The lack of generalism across the groups is remarkable, and it contrasts with results suggesting that most abundant Amazonian rainforest trees are generalists (Pitman et al. 2001, 2013, Arellano et al. 2014). The untransformed base cation gradient in the Melastomataceae study (**I**) ranged from 0.07 to 33.59 cmol(+)kg^{-1}. The corresponding gradient in the fern study (**II**) was longer, from 0.02 to 54.36 cmol(+)kg^{-1}. On average, the estimated relative width of the cation niches of species on the log10 -transformed gradient was approximately a quarter of the gradient (24%) in ferns and approximately a third (31%) in Melastomataceae. These relative lengths are not directly comparable between ferns and Melastomataceae, since they were estimated from gradients that differ in length, and with different methods: weighted averaging for ferns and HOF modelling for Melastomataceae. It is nonetheless clear that most species in both groups had cation niches that spanned only a relatively small part of the gradient. This suggests a similar pattern in ferns and Melastomataceae that Phillips et al. (2003) found for trees, where 75% of species were significantly related to edaphic habitat. Fine et al. (2005) reported no soil generalists among 35 Burseraceae tree species.

Community-level response to the base cation gradient has often been reported for Amazonian trees, ferns, and palms (Phillips et al. 2003, Guèze et al. 2013, Zuquim et al. 2014, Cámara-Leret et al. 2017). The strong species-level responses of ferns and Melastomataceae in papers **I** and **II** are in line with the earlier community-level results, and the pervasive lack of species-level generalism suggests that the community-level patterns are not solely created by a few abundant, edaphically specialised species, but are a property of the majority of species in at least ferns and Melastomataceae. In contrast to this, species-level results on rainforest trees in Bolivia (Toledo et al. 2012) and Panama (Condit et al. 2013) suggest that less than 50% of the species have a detectable response to soil chemistry, and in those studies, soil was less significant than climate as a driver of community assembly. One possible reason for this discrepancy is that the climatic gradient between central and northern Bolivia, and between the east and west coasts of the Panama isthmus, is more pronounced than that in papers **I** and **II**, which both comprised only areas within lowland Amazonia. The climatic effect might have reduced the apparent effect that soils have on species distributions, if species distributions are constrained by climate, and soil variation within each climatic zone is smaller than it was in our study. A second possible reason is that the number of individuals per species in tree studies is often much smaller than in fern and Melastomataceae studies, which reduces the likelihood of detecting statistically significant niche specificity. It is also possible that ferns and Melastomataceae are truly different in this respect from trees. At least in ferns, dispersal limitation is not likely to be as strong as in trees (Zuquim et al. 2012). However, Amazonian palms behave like ferns and Melastomataceae did in papers **I** and **II**: community-level response is reflected in species-level responses to soil cations (Cámara-Leret et al. 2017). The difference between trees and other rainforest plants might also be explained by the fact that trees tend to have wider niches than understorey plants (Ruokolainen & Vormisto 2000, Duque et al. 2002, Chust et al. 2006), but I did not find any such pattern between high- and low-statured Melastomataceae (**I**).

Most Melastomataceae (70%) had their optima in the lower half of the base cation gradient (paper **I**). Their niches were also slightly narrower in that part of the gradient; species with high optima had a weak but significant tendency to have broader niches. This is in accordance with the theory on interspecific competition, which states that intense competition leads to narrow niches (McNaughton & Wolf 1970). Such narrowing of niches happens, e.g., in Finnish forest plants (Heikkinen & Mäkipää 2010). On the other hand, Amazonian Melastomataceae form only a small part of the understorey flora, and they do not necessarily compete principally with each other, but with other plant groups within the same guild. This renders the explanation involving interspecific competition unlikely. It has been observed earlier that Melastomataceae species richness tends to be highest in the lower half of the base cation gradient (Tuomisto & Ruokolainen 2005). This can be a case of niche conservatism, but it is not exclusive, since there are also Melastomataceae whose niches are restricted to the high-cation part of the gradient. Soil niche conservatism within the Amazonian lowland rainforests is remarkable, since not very long ago the forests were thought to be environmentally rather uniform. Phylogenetic soil niche conservatism has been previously reported for some Amazonian fern and palm genera (Lehtonen et al. 2015, Cámara-Leret et al. 2017), and for trees in Panama (Condit et al. 2013).

In contrast to Melastomataceae, the optima of *Adiantum* and *Lindsaea* species (both genera combined) were relatively evenly spread along the base cation gradient (paper **II**). An earlier comparable example of this is the even spread of base cation optima in *Polybotrya* ferns (Tuomisto 2006). Amazonian palms seem to have a contrasting pattern: more species had their base cation optima close to either end of the gradient than in the middle section (Cámara-Leret et al. 2017). In *Adiantum* and *Lindsaea,* marginal niches were on average narrower than intermediate niches. This is often interpreted as being due to niche truncation at the gradient ends, which can indeed be the case particularly in the lower end of the gradient, since the nutrient-poor white-sand forests were poorly represented in our dataset. High-nutrient sites, on the other hand, were better represented, and it is unlikely that they would be relatively more common in nature than they were in our sample.

Most *Adiantum* had their optimum in the high-cation part of the gradient, but there were some *Adiantum* species in all but the poorest soils. All species of *Lindsaea* occurred predominantly in poor soils and were absent from the richest soils. The highest species richness of *Adiantum* and *Lindsaea* ferns was, therefore, in the intermediate-low part of the cation gradient, where both genera had many species. If we look at central Amazonia separately, however, the species richness of *Adiantum* and *Lindsaea* increased linearly with increasing soil cation concentrations. This is likely because central Amazonia has, on average, poorer soils than the two Andean regions.

The unimodal response type was universally common. Using abundance data, 84% of the *Adiantum* and *Lindsaea* species and 74% of Melastomataceae showed a unimodal response to the base cation gradient. This meets the theoretical expectations of unimodal responses on long gradients (Gauch & Whittaker 1972, Austin & Smith 1989, Oksanen & Minchin 2002). Toledo et al. (2012) modelled the species responses of trees in Bolivian Amazonia and found that unimodal responses were rare (only 6% of the soil chemistry responses). They found the result surprising, since the soil gradient was rather long (corresponding to a 70-fold change in Ca and Mg concentrations). The gradient studied in paper **II** is even longer than that (more than 2500-fold), which may partly explain the difference in the percentage of unimodal response shapes between this study and Toledo et al. (2012). There are methodological differences as well. In paper **II**, the modelling method was HOF modelling, whereas Toledo et al. (2012) used multiple logistic regression.

Most of the unimodal responses were skewed. In ferns, using presence-absence data led to the symmetrical unimodal response becoming more common at the expense of the skewed unimodal response, and also non-unimodal responses, particularly the plateau model, becoming more common. Since these models have fewer parameters than the skewed unimodal model, it can be said that using presence-absence data led to selecting simpler models. Abundance data produces more nuanced models, but on the other hand, the effect that single large abundances (outliers) have on the model shapes was large in several cases. This was particularly true with some *Adiantum* species that are capable of spreading vegetatively with a horizontal rhizome, and thus attaining very high local abundances at some, but not all, sites where they are present. Presence-absence data and abundance data both have advantages and disadvantages in species response modelling: presence-absence is less sensitive to outliers, but abundance data can at least theoretically better distinguish sites where a species is truly able to maintain a viable population.

Frequent species tended to have a skewed unimodal response in both ferns and Melastomataceae, and the other model types were more often associated with infrequent species. It is possible that some of the other species' response shapes would also have become asymmetric unimodal instead of a plateau shape or symmetric unimodal shape, should the number of observations still increase. The skewed unimodal model's demand for more occurrences is partly explained by the HOF model selection rules: the skewed unimodal model needs one parameter more than the symmetric unimodal model does, and therefore it is not likely to be chosen unless robustly backed up by the distribution of a large number of occurrences along the gradient (Michaelis & Diekmann 2018).

According to the asymmetric abiotic stress hypothesis, species response shapes on gradients are skewed because competition intensifies in the most favourable part of the gradient, and this is the direction of the skew (Austin 1990, Austin & Gaywood 1994, Rydgren et al. 2003, Normand et al. 2009). It is often interpreted that the optimal conditions, associated with intense biotic competition, are in the middle of the gradient, and physiological stress becomes more intense towards the gradient margins. In *Adiantum* and *Lindsaea*, it cannot be assumed that the optimal conditions for all species would be in the middle of the gradient, but the optimal habitat is likely different between species of *Adiantum* (mostly the high-cation end) and *Lindsaea* (mostly the low-cation end; Zuquim et al. 2014). Many *Lindsaea* responses were truncated at left and *Adiantum* responses at right, leaving the tail towards the unfavourable part of the cation gradient for each genus, respectively—a result that appears at odds with the asymmetric abiotic stress hypothesis. Based on this data, it is not possible to say whether it is competition or physiological intolerance that prevents these fern species from growing in all parts of the base cation gradient. Experimental comparisons of niche breadth in field conditions, such as transplant experiments, could help in clarifying the mechanisms.

3.2 Patterns of commonness and rarity

The various aspects of commonness include frequency, local abundance, geographical range size and niche width. These tend to be intercorrelated. The niche breadth hypothesis (McNaughton & Wolf 1970, Hanski 1982, Brown et al. 1984, Slatyer et al. 2013) suggests that locally abundant species are also regionally frequent, and that common and geographically widespread species tend to have wider niches than rare and geographically restricted species. There are earlier results on Amazonian plants that conform to these expectations (Pitman et al. 2001, 2013, Emilio et al. 2010, Arellano et al. 2014).

My results on *Adiantum* and *Lindsaea* mostly fail to fill these expectations, since geographical range was the only aspect of commonness that was significantly correlated with niche width (paper **II**). Minimum niche width did increase with geographical range size, but some of the most geographically restricted species had niche widths that corresponded to the average niche width of the most widespread species (paper **II**, Fig. 5). In Melastomataceae (paper **I**), the correlations between niche width and aspects of commonness were likewise weak and may reflect little more than the fact that many observations are likely to span over a larger range on the gradient than few observations (Table 2).

Local abundance was not correlated with either frequency or geographical range size in *Adiantum* and *Lindsaea*, nor was it correlated with frequency in Melastomataceae. There was a positive correlation between frequency and geographical range size, mainly because geographically restricted species did not attain very high frequencies. A notable exception was *Lindsaea bolivarensis*, which was geographically restricted to northwestern Amazonia but still attained a relatively high frequency. Several geographically widespread species were infrequent; a large range did not guarantee a high frequency.

The geographical distribution of the preferred habitat is probably a strong determinant of species distributions (Brown 1984). However, in my study, local abundances of individual species varied widely even when base cation concentrations did not. This probably reflects variation in base cation concentration within transects, and likely also the influence of other environmental variables, and biotic interactions. It is also true that evolutionary history and random processes influence these patterns. Populations of the same species can become ecologically distinct in different areas. Taking evolutionary time scale into account would enable deeper understanding of niches (Lehtonen et al. 2021). However, that would have necessitated producing phylogenies of Melastomataceae and the two fern genera, which is beyond the scope of this thesis. I treat species as static, separate entities that possess certain ecological characteristics, and I look at the species distributions as they can be observed presently. Also indicator species systems in Europe handle species in the same way.

Overall, my results do not provide strong support to the niche breadth hypothesis, even though the assumptions of the hypothesis (Brown 1984) are to a large extent met by the studied species: they are closely related (ferns within genus, Melastomataceae within family), and ecologically so similar that the species are likely to differ in only a few niche dimensions. The correlation between local abundance and frequency was likewise weak in a study on trees in Bolivian

Amazonia (Toledo et al. 2012). The fact that infrequent species with a narrow niche attained high local abundances is consistent with the idea of trade-offs in species tolerances leading to niche differences (Silvertown 2004), but not with the niche breadth hypothesis (Brown 1984, Slatyer et al. 2013). More research on the relationships between the various aspects of commonness among Amazonian plants is obviously needed. In the next few paragraphs, I will compare my results on patterns of commonness and rarity with some of the comparable earlier studies on Amazonian plants.

Table 2. Correlations of four different aspects of commonness with niche width in the studied species. Frequency is the number of occupied transects. Total abundance is the total number of individuals in the data. Local abundance is the average number of individuals per area in those transects where the species was registered (average density when present). Geographical range size is the longest distance between any two transects where the species was registered (not estimated for Melastomataceae due to the smaller regional extent of sampling).

ASPECT OF COMMONNESS	***ADIANTUM* AND *LINDSAEA***	**MELASTOMATACEAE**
FREQUENCY	Not significant	Weakly positive
TOTAL ABUNDANCE	Not significant	Not significant
LOCAL ABUNDANCE	Not significant	Weakly negative
GEOGRAPHICAL RANGE SIZE	Positive	

Pitman et al. (2001) found that an oligarchy of approximately 150 tree species dominates Amazonian forests close to the Andes in southern Peru and eastern Ecuador. They concluded that the oligarchic species, which are frequent and abundant over wide areas, have wide environmental niches and are therefore able to occur in a high proportion of the available area. It is interesting to compare the most common species in that study, the palm *Iriartea deltoidea*, with the most common species in paper **II**, *Adiantum humile*. Cámara-Leret et al. (2017) made a HOF model for the abundance of *I. deltoidea* on the base cation gradient, and it is surprisingly similar to the one I made for *A. humile.* These two species do have relatively wide base cation niches, but HOF models predict that both will be abundant on relatively rich soils only. The way Pitman et al. (2001) calculated local abundance differs from that in my study: they calculated it by dividing the total abundance by the number of sampling units, whereas in papers **I** and **II**, local abundance was recorded for each sampling unit separately. Pitman et al. (2001) might have missed variation in local abundance that is actually there, leading to the conclusion that *I. deltoidea* occurs without habitat preference in western Amazonia. It is also true that the soil gradient in western Amazonia is much shorter than the gradient in papers **I** and **II**, but even if only the northwestern and southwestern subsets of my data are taken into account, *A. humile* occurs in a similar percentage of sampling units (50–60%) as *Iriartea* did in Pitman et al. (2001).

Iriartea deltoidea belongs to the suite of oligarchic species and occurs in all major forest types of the lowland rainforests in northern Peru (Pitman et al. 2001, Kristiansen et al. 2012) and is also part of the oligarchy in Bolivia (Arellano et al. 2014). In paper **IV**, I studied its abundance and found that it was abundant on the Pebas Formation but almost completely absent from the Nauta Formation. This contrasting result might be due to at least two different reasons. The first is related to sampling strategy: I only counted individuals taller than 1 m, in order to avoid confusion with juveniles of other, related palm species. It is possible that the Nauta Formation does have *I. deltoidea* individuals smaller than 1 m that fail to reach larger sizes due to environmental constraints. In that case, Nauta Formation forests would harbour sink populations of this species, which has more recently been modelled to reach high abundances only in forests that are more nutrient-rich than most sites on the Nauta Formation (Cámara-Leret et al. 2017). Secondly, *I. deltoidea* is a species used in house construction. In this study area, the Nauta Formation is closer to the riverside villages than the Pebas Formation, raising the possibility that the better accessibility of the Nauta Formation has led to more selective logging of *I. deltoidea* there. Based on this data, it is not possible to prefer either of these explanations. Both of them might be working simultaneously to produce the observed distribution pattern.

Pitman et al. (2001) also suggested that rare species might be specialists, i.e., have narrow environmental niches. The niches of rare trees are very difficult to quantify due to their low densities in the landscape. In light of our results on ferns and Melastomataceae (**I**, **II**), some rare species do have narrow niches, but some also have broad ones. It has to be kept in mind, however, that even if a species has a broad niche along the cation gradient, it might have a narrow niche along some other relevant environmental gradients. In ferns, a likely candidate is soil moisture, which is often controlled by local drainage conditions (Moulatlet et al. 2014). A species might be indifferent to soil base cation concentration as long as its moisture niche, e.g., a flood-prone creek bank, is available, and it attains high densities there irrespective of the base cation concentration. It may also be that the flooding causes higher base cation concentrations on these microsites, and this kind of small-scale variation is not captured by the soil sampling design used in my studies (Damasco et al. 2013).

Arellano et al. (2014) found that oligarchic tree species show broader niches than other species. They argue that the result strongly supports the niche breadth hypothesis. In their study, niche breadth was positively correlated with both frequency and local abundance. Although their results seem convincing and contradictory to ours, there are a couple of differences between these studies that make direct comparisons difficult. Arellano et al. (2014) use an index of relative generalism, where the distribution of a species is compared to the distribution of the environmental variable. According to that index, those species that are distributed in the landscape in the same way as the resource, are generalists. A possible misinterpretation is that if a species is specialised to the most abundant kind of resource in the landscape, it will appear more generalist than a species that is specialised to a rare resource. For example, in the fern data I used, one of the most abundant species is *Adiantum tomentosum*, a fern that has a relatively narrow niche. It is specialised to the most widely available section of the cation concentration gradient in the dataset. Using an index of relative generalism, it would probably appear to be both an abundant species and a generalist, leading to a conclusion supporting the niche breadth hypothesis. Another difference between the studies is that the results of Arellano et al. (2014) are based on samples within an area of 200 km x 200 km in southwestern Amazonia. The soil gradient covered is probably shorter than in my studies, which makes it more likely to find both oligarchic and generalist species.

Adiantum and *Lindsaea* had species in all rarity categories coined by Rabinowitz (1981), except in the category defined by the combination “small geographical range, wide niche and low local abundance”. The category “small geographical range, wide niche and high local abundance” had only a single species (Table 3). Rabinowitz (1981) predicted that these categories are likely to have few species.

A clear majority of *Adiantum* and *Lindsaea* were large-range species. Pitman et al. (1999) classified southwestern Amazonian tree species into Rabinowitz categories. They used occurrence outside of the department of Madre de Dios in Peru (approximately 78 000 km^2), the administrative region of their plots, as a limiting criterion between a small and a large geographical range. They found that all of the species occurred outside of Madre de Dios, i.e., all were large-range species. Compared to that, some of the ferns studied here seem to have a more restricted range, but large ranges also predominate in *Adiantum* and *Lindsaea*. On the other hand, both study regions of Pitman et al. (1999), Madre de Dios in southern Peru and Yasuní in Ecuador, represent the higher end of the soil base cation gradient of lowland Amazonia. Geographical ranges of trees in that study can, therefore, also be smaller than those of *Adiantum* and *Lindsaea* ferns, if the trees eventually do not occur outside the rich-soil regions of western Amazonia. For many of the western Amazonian oligarchic species, this is indeed the case (Draper et al. 2021).

There were both frequent and infrequent *Adiantum* and *Lindsaea* species that had a large geographical range (paper **II**), but nearly all species with a restricted geographical range were infrequent, which led to a high correlation between frequency and geographical range (Pearson correlation 0.70, $p<0.0001$). A comparable result was reported by Arellano et al. (2014), who found both broad- and narrow-niche oligarchic species. In this case, the infrequent but geographically widespread species would correspond to the narrow-niche oligarchs. According to Fine et al. (2005, 2010), the same species of trees inhabit islands of white-sand soils over large expanses of Amazonia. This kind of edaphic specialisation combined with a large geographical range seems also likely in the geographically widespread but infrequent *Adiantum* and *Lindsaea*. If the edaphic habitat is infrequent but geographically widespread, the taxa specialised to it will show the same pattern, at least in the absence of strong dispersal limitation.

Four species belonged to the most restrictive rarity category (small geographical range, low local abundance and narrow niche). Of these, *L. hemiglossa* and *L. schomburgkii* are white-sand specialists, and that soil type was not well represented in our dataset.

Based on this data, it is possible to draw conclusions about relative abundance and rarity. Even though this is certainly the largest Amazonian fern dataset so far, its spatial coverage is still very sparse compared to that of vegetation surveys in temperate regions, such as northern Europe (Hedwall et al. 2019) or eastern Asia (Choe et al. 2019). Therefore, caution is needed with conclusions about absolute rarity.

New data is most likely to affect the estimated geographical range sizes, which can dramatically grow if a species is found in a new area. However, niche position and niche width seem to be much more consistent and are not likely to change much with new data, unless species limits change as well: a single, widespread species might turn out to represent several hitherto unrecognised cryptic species, which are geographically restricted. This happened with the oligarchic tree species *Protium heptaphyllum* (Damasco et al. 2021), and with the widespread Amazonian fern *Metaxya rostrata* (Cárdenas et al. 2016). Of the species I studied, *Adiantum humile* is a good example of a likely candidate for a comparable change in species limits (H. Tuomisto, personal communication).

Besides such complex species, changes in niche position and niche width with new data are most probable in those species that have their optima close to either end of the gradient. Truncated responses can become unimodal if new samples are found in areas with very poor or very rich soils. However, real-life gradients are not infinite, and even though soils poorer than these (such as the white sands) may be relatively frequent, it is unlikely that richer soils than these would be very common in Amazonia. Those are most likely to be found close to the Andes, where volcanism enhances soil fertility.

Table 3. Amazonian *Adiantum* and *Lindsaea* species in the eight rarity categories of Rabinowitz (1981), based on 1061 presence-only transects spread over three regions: northwestern, southwestern and central Amazonia. Local abundance is based on a subset of 402 abundance-quality transects. The category in the upper left corner is the "not rare" category. Three representative species are shown for each "large geographical range" category. The "small range" categories show all species. Category limits are the dataset median for local abundance, 1000 km between the two most distant observations for geographical range, and 25% of gradient length for niche width.

	LARGE GEOGRAPHICAL RANGE		**SMALL GEOGRAPHICAL RANGE**	
LOCALLY ABUNDANT	**10 species** e.g. *A. humile* *A. pulverulentum* *L. divaricata*	**8 species** e.g. *A. tomentosum* *A. cinnamomeum* *L. lancea*	**1 species** *A. poeppigianum*	**3 species** *A. pseudocajennense* *A. wilsonii* *L. bolivarensis*
LOCALLY SCARCE	**6 species** e.g. *A. nodosum* *L. phassa* *L. digitata*	**11 species** e.g. *A. tuomistoanum* *A. diogoanum* *L. guianensis*		**4 species** *A. fuliginosum* *L. hemiglossa* *L. coarctata* *L. schomburgkii*
	Wide niche	**Narrow niche**	**Wide niche**	**Narrow niche**

3.3 Consistencies and differences in species responses

In paper **I**, I compared the optima and niche widths of Melastomataceae between two regions (NW vs. SW + C Amazonia). In paper **II**, I compared the *Adiantum* and *Lindsaea* species optima and niche widths obtained with different modelling methods (HOF and WA), data types (abundance and presence-absence) and based on three different regions (southwestern, northwestern, and central Amazonia).

For Melastomataceae, the optima were relatively consistent between regions, but the degree of consistency was dependent on species frequency: the optima of frequent species were more consistent than those of infrequent ones (paper **I**). A similar trend was observed in the HOF models for ferns: infrequent species often had a different HOF model based on abundance data than based on presence-absence data. The results are more tentative for infrequent species and more robust for species with many observations.

Although the optima tended to be consistent, niche widths from different regions were not significantly correlated in Melastomataceae. One reason for this might be that I estimated Melastomataceae niche widths with HOF modelling only, and not with weighted averaging. HOF models produce more variable niche widths than weighted averaging does, and even the selected HOF model type might differ between the regions.

The species optima of *Adiantum* and *Lindsaea* were consistent across modelling methods and data types (paper **II**). The correlations between the optima from NW and SW Amazonia were high. The optima from central Amazonia differed from either of the western regions. This was because the optima estimated based on central Amazonian data had a relatively narrow spread: the lowest species optima tended to be higher, and the highest species optima tended to be lower than on the basis of western Amazonian data.

The lower optima of rich-soil species in central Amazonia can be due to the poorer availability of rich-soil habitat than in the western regions; several rich-soil species were also absent from central Amazonia. Local adaptation is also a possible explanation, but this remains to be tested. It could occur in a similar manner as in temperate Europe, where rich-fen plant species have adapted to acidic, low-nutrient conditions in glacial refugia in the Balkans (Hájková et al. 2008). In those refugia, mostly only acidic mires were available, a situation reminiscent of the generally cation-poor central Amazonia.

The higher optima of poor-soil species in central Amazonia can be due to niche shifts in a different climate. The more seasonal rainfall pattern of central Amazonia leads to longer dry seasons, and drought is a stress factor for most ferns. The suboptimal climatic conditions in central Amazonia might induce niche shifts towards higher cation concentrations in *Adiantum* and *Lindsaea*, in a similar way that a cold climate and a short growing season cause some neutrophilic temperate European plant species to retreat from acidic sites at higher latitudes (Reinecke et al. 2016).

3.4 Implications for indicator species use

In paper **I**, I modelled the responses of individual Melastomataceae species to the gradient of base cations in the soil, and in paper **II**, I did the same for *Adiantum* and *Lindsaea* ferns. In paper **III**, I studied prediction of different soil variables using all species in the Melastomataceae family.

The results show that Melastomataceae species composition can be used successfully to predict the concentrations of calcium, magnesium, potassium and the sum of base cations in the soil at Amazonian lowland rainforest sites. With other soil properties, such as phosphorus concentration, aluminium concentration and loss-on-ignition, the prediction accuracy is suboptimal, or even poor. The k-NN and weighted averaging methods performed approximately equally well, so for prediction purposes, the choice of method is not of great importance. The use of plants in bioindication facilitates surveys since, as the surveyor accumulates experience, plant identification gets faster and the need to collect voucher specimens decreases. This does not happen to soil sample collection and analysis: each sample is equally costly in time and money. The concentration of base cations is positively correlated with primary productivity, which makes the mapping useful for land-use planning. Sadly, agricultural areas are often made in primary rainforests that have too infertile soils, and the forest gets sacrificed for next to no agricultural gain. In this context, rapid and cost-efficient soil mapping using indicator species could be of help. Melastomataceae species composition also covaries with other plant groups (Ruokolainen et al. 2007) and can therefore be used as a proxy for general floristic variation in e.g., assessing the representativeness of conservation area networks. Indicator values for 260 Melastomataceae species and morphospecies for the concentrations of total base cations, magnesium, calcium and potassium are found in paper **III**, and base cation niche characteristics of 81 Melastomataceae identified to a named species are found in paper **I**.

The *Adiantum* and *Lindsaea* species niches on the cation gradient suggest that they also form a promising indicator group. Together, their niches cover the sampled cation gradient in such a way that each gradient quarter has many species (paper **II**). There are eight *Adiantum* and *Lindsaea* species that meet the criteria for good indicators according to Rabinowitz's (1981) rarity classification: they are geographically widespread and locally abundant but have a relatively narrow soil niche (Table 3). Also the species in the category termed "not rare" are potential indicators, since despite their commonness, they are not cation generalists. They do have a relatively wide niche, but with information on their abundance, one can use them for bioindication. In a similar way, the Finnish forest site type system uses cover percentages of common Ericaceae dwarf shrubs to separate broad forest site types (Kuusipalo 1985), even though these species have very broad and overlapping tolerances (Heikkinen & Mäkipää 2010). The site type classification is then completed with narrow-niche herb species to distinguish subclasses of the broader forest site types. A comparable practice might be feasible in Amazonia, using *Adiantum* and *Lindsaea*. Their advantages include the recognisable overall appearance that makes these genera easy to distinguish from other ferns. Of course, the number of species in them is considerably smaller than the number of all fern species, which makes inventories much quicker.

There are many *Adiantum*, *Lindsaea* and Melastomataceae species with a geographic distribution that spans two or more Amazonian regions (**I, II**). Despite their wide distribution, some of them have quite narrow niches: they occur throughout Amazonia, but are present only on soils that belong to a certain section of the soil cation concentration gradient. This leads to the conclusion that a rich-soil community or a poor-soil community of ferns or Melastomataceae might be recognisable anywhere in Amazonian lowland rainforests. That situation is quite different from the communities of understorey tree species, whose floristic composition differs between Amazonian regions (Draper et al. 2021).

Which indicator group should one then choose for the concentration of base cations in Amazonian soils, Melastomataceae or *Adiantum* and *Lindsaea*? A firm conclusion would necessitate further research, but some similarities and differences between these groups can be listed here. Both are easy to separate from other plants based on their appearance. Both are also relatively easy to collect, since all *Adiantum* and *Lindsaea* and most Melastomataceae grow at convenient heights, and do not have very large leaves. Melastomataceae has a higher species richness, and due to the many unnamed morphospecies, more difficult taxonomy as well. With Melastomataceae I would opt to work with a subset of species that are both relatively common and easy to identify (paper **III**). The optima of Melastomataceae are concentrated in the lower half of the soil cation gradient, whereas those of *Adiantum* and *Lindsaea* are more evenly spread. In terms of average niche width, the two groups are probably roughly similar.

3.5 Structure of lowland rainforest on two geological formations

The *terra firme* rainforests on the Nauta and Pebas formations are known to have clearly different floristic compositions and soil properties. The floristic difference, measured as the community composition of the most abundant ferns, was clear in this study, but the forest structural differences proved elusive (paper **IV**).

During the fieldwork, I developed a personal notion of the two forest types, based on the overall colour of the understorey. Nauta Formation appeared “brown forest”, and Pebas Formation “green forest”, probably due to the higher abundance of large-leaved understorey herbs in the latter. For pole-sized trees, there was a significant difference in stem density (Nauta Formation was more dense, as expected), but for the density of other size classes, and canopy openness, the results were inconclusive. The power tests concluded that, assuming constant variance and a difference in means of similar size, the number of field transects should have been 3–10 times larger in order to detect these differences as statistically significant. Certainly, one cannot know what would have happened to the variance and the difference in means with a larger sample size. The conclusion is that even if a structural difference might exist, it is not easily measured on the ground.

Comparison with the reference transects on the Landsat image showed that the field transects might have underestimated the regional differences between the two forest types. Moreover, the result on sapling density based on reference transects was opposite to that based on field transects: the Pebas Formation forests appeared to have much higher sapling density. The result based on the field transects had been in the other direction, albeit not significantly so.

Regression modelling with Landsat and fern data produced a much more robust model than the modelling with Landsat and forest structural data did. This happened despite the fact that the fern community was not sampled in its entirety, but the fern data consisted of only the three most abundant species per observation point. This further highlights the difficulty of detecting structural differences between forest types that do differ floristically.

Stem density does not necessarily always have a simple linear negative correlation with understorey light availability. White sand forests of Amazonia are known to combine high stem density with high light in the understorey (Anderson 1981), and comparable results have also been obtained in Bornean dipterocarp rainforests (Russo et al. 2012). In this study, canopy openness was negatively correlated with the density of large stems, but not with that of saplings and poles.

Earlier studies had produced similar results of invariable structure of tropical rainforests growing on different geological surfaces (Gentry 1982). Also, some studies had found that more nutrient rich forests were more open (Slik et al. 2010), and others the other way around (DeWalt & Chave 2004, Paoli et al. 2008). Baraloto et al. (2011) concluded that the density of large stems is higher in nutrient-rich forests, whereas that of small stems is higher in nutrient-poor forests.

Almost simultaneously with the publication of this study, a study on structural differences between the Nauta and Pebas formations in northern Peru, using field data of ferns and soils combined with airborne spectroscopy and LiDAR data collected with the Carnegie Airborne Observatory (CAO), was published (Higgins et al. 2015). The study found marked differences between the geological formations in the vertical distribution of the vegetation, canopy reflectance, and gap frequency. Pebas Formation forests had, on average, a lower, more variable, and more open canopy than Nauta Formation forests did. Gaps were more frequent on the more nutrient-rich Pebas formation. Those results are in accordance with earlier studies indicating less dynamic forests on the nutrient-poor central Amazonian soils compared to those in western Amazonia (Phillips et al. 2004, Malhi et al. 2006). They also correspond well to the Brazilian terminology of closed and open forests.

The only significant forest structural difference found in my study, the higher density of pole-sized trees on the Nauta formation, appears to contradict the results of Higgins et al. (2015). Trees in that trunk diameter size-class in western Amazonia are on average 6–12 m high (Feldpausch et al. 2011), and their canopy should comprise some metres below that height. The results of Higgins et al. (2015) show that Pebas Formation forests have more vegetation at those heights than the Nauta Formation forests do. This difference can, however, be due to lianas, palms, or large herbs, which can have leaves at those heights. The LiDAR instrument detects all vegetation, but I only measured trees, so I could not detect variation in these groups. The overall conclusion is that even though the differences in forest structure between the formations proved difficult to measure on the ground, they still clearly do exist.

Even more recently, Doughty et al. (2022) used GEDI, a LiDAR instrument based on the International Space Station, to estimate forest canopy structure across all three major tropical rainforest areas (Amazonia, central Africa and southeast Asia). They detected a change in canopy structure that coincided with the boundary of the Pebas Formation with the less nutrient-rich forests in central Amazonia. The Pebas Formation was interpreted to have a more stratified and complex structure than the central Amazonian forests, which correspond to the Nauta Formation at least at their western margin. According to Doughty et al. (2022), Nauta Formation forests have a single vertical peak in vegetation volume, whereas Pebas Formation forests have two peaks. Doughty et al. (2022) drew the conclusion that one-peak forests have a more open canopy, which seems contradictory to the Brazilian map terminology. However, it is also possible that the more stratified Pebas Formation forests appear more heterogeneous to the human observer than the single-peak Nauta Formation forests do. This would have led the observers that made the Brazilian classification to conclude that the Nauta Formation forests are denser.

3.6 Drivers of epiphyte abundance in cloud forests

In paper **V**, I studied the drivers of epiphyte abundance in the Andean pre-montane cloud forests. Canopy openness was the most important forest structural variable affecting epiphyte abundance. Microclimatic variables were correlated with canopy openness, and therefore their inclusion did not improve the models predicting epiphyte abundance. The responses of different epiphyte groups to forest structure and microclimate were varied.

The bryophyte cover decreased significantly with increasing maximum temperature but did not respond significantly to minimum humidity. Bryophyte cover had a positive relationship with tree height and a negative one with canopy openness. Basal area did not have a significant effect on bryophyte cover. Model averaging resulted in the selection of maximum temperature as the most important variable explaining bryophyte cover.

The different vascular epiphyte groups also showed significant responses to both microclimatic and forest structural variables. There were not enough records of orchids and palms to make conclusions about their responses, so I left them out of the analysis. The scarcity of orchids in this sample was most likely due to the fact that I only sampled the tree trunks. Epiphytic orchids typically grow higher than this, on the canopy branches (Krömer et al. 2007, Petter et al. 2016). I only report results on aroids, bromeliads and ferns. All of them showed a significant positive relationship with minimum humidity. Ferns and aroids also showed a significant negative relationship with maximum temperature. The responses of the three vascular epiphyte groups to forest structural variables were varied. The abundance of aroids increased significantly with basal area and tree height, whereas for ferns this relationship was weaker, and it was nonsignificant for bromeliads. The abundance of all three groups was negatively correlated with canopy openness. The selection of the most important explanatory variables with model averaging resulted in basal area being most important for aroids, and canopy openness for ferns and bromeliads. Minimum humidity was selected as an important variable for the pooled abundance of all three vascular epiphyte groups.

The results are in overall agreement with earlier studies suggesting that microclimate plays an important role in controlling the abundance of different epiphyte groups (Nadkarni 2000, Holz & Gradstein 2005, Gradstein 2008, Nöske et al. 2008). The most important finding was the importance of canopy openness to all epiphyte groups, and its correlation with microclimate. Adding microclimatic variables to the models explaining epiphyte abundance did not improve model performance, and I interpret this as an indication of a tight coupling between forest structure and microclimate. For epiphyte studies, this is good news since canopy openness is much more inexpensive and rapid to measure than the microclimatic variables are.

The fact that bryophyte cover declined with increasing maxT is attributable to the increased water loss from bryophytes in hot conditions (Zotz et al. 1997). Bryophytes also have a lower chlorophyll concentration than vascular plants, and high temperatures are more likely to result in a negative carbon balance in bryophytes than in vascular plants, due to nighttime respiratory losses exceeding daytime photosynthetic gains (Frahm 1990, Zotz et al. 1997, Martin & Adamson 2001). In our study, vascular epiphyte abundance was a better proxy for air humidity than bryophyte cover. In some earlier studies, bryophyte cover was a good proxy for air humidity in cloud forests (Wolf 1993, Gehrig-Downie et al. 2011, Karger et al. 2012). However, those studies did not compare bryophytes with vascular epiphytes, and they used average values of humidity, whereas we used minimum humidity, so the results are not entirely comparable. An ecological explanation for our result is that the sensitivity of young individuals of vascular epiphytes to periodic drought makes these plants more sensitive to low air humidity than the more desiccation-tolerant bryophytes (Schmidt et al. 2001). The colonisation of bare branches in the cloud forest canopy—an environment with periodic drought—is pioneered by bryophytes, not vascular epiphytes (Nadkarni 2000).

Another possible explanation is related to our field sampling and the altitude of the plots where we collected microclimate data. Our field sampling was done in the dry season (June–August) of two successive years (2014 and 2015), and in the first year the plots were at lower altitudes than in the second year (the average altitude of the first year plots is 1271 m, and for the second year plots it is 1362 m). The microclimate data logging took place between the two sampling efforts, in the rainy season and transition to the dry season (January–May 2015), and the data loggers were located in the first-year plots only. This was partly due to problems getting a permit to work inside the Alto Mayo protected forest in the first sampling year. If the microclimate had also been measured in the plots of the second year, at a higher mean altitude, the relationship between microclimate and bryophyte cover would probably have been stronger, since bryophyte cover in tropical mountains usually increases with elevation (Frahm & Gradstein 1991). In our data, this is indicated by the bryophyte cover being consistently higher in the second-year plots, even though there is no difference in canopy openness, tree height, or basal area between them and the first-year plots (paper **V**, Fig. 2).

The significant negative relationship of bryophyte cover and canopy openness became nonsignificant when the second year plots were included in the analysis. However, with tree height a significant positive relationship was maintained even when the second year plots were included. The conclusion is that since the second year plots were on average at ca. 100 m higher altitude, they probably also had a lower maximum temperature and a higher minimum humidity, because potential evapotranspiration decreases by 0.1 mm/d per 100-m increase in elevation (Thornthwaite 1948). The second year plots did not have a lower canopy openness than the first year plots, but they did have a higher bryophyte cover, likely because of the more amenable microclimate for bryophytes at higher altitude (irrespective of canopy openness).

A high basal area and high tree height were more important for aroids than for the other vascular epiphyte groups. The ecology of aroids offers an explanation: many adult aroids are comparatively heavy plants; many of them are also hemiepiphytes, i.e. they either start their life cycle on the ground, or start as epiphytes and reach the ground later with adventitious roots (Mantovani 1999, Krömer et al. 2007). This leads to them growing at lower positions on the trunk than most other epiphytes, and explains why they benefit from larger trees (Petter et al. 2016).

The abundance of all vascular epiphyte groups and bryophyte cover was negatively correlated with canopy openness. I interpret this as canopy openness being connected to microclimate because, in more open forests, drier and sunnier conditions prevail. Caution is needed, however, because 1) Our microclimatic measurements did not cover the whole year and 2) The plots where microclimate data was collected were on average at an approximately 100 m lower altitude than the rest of the plots, and in more fragmented surroundings.

In addition to microclimate, the presence of large trees is another possible explanation for the variation in epiphyte abundance. The presence of large trees leads to lower canopy openness, and large trees offer more attachment sites to epiphytes than small trees do (Nöske et al. 2008). The presence of large trees, however, should also lead to a higher basal area, but basal area did not have a significant effect on most of the epiphyte groups. Canopy openness and basal area were negatively correlated in our data, so it is still possible that these two forest structural variables both affected epiphyte abundance. A more open canopy makes the microclimate warmer and drier, and high canopy openness often also means that there are fewer large trees available for epiphyte attachment.

Selective logging of the pre-montane cloud forests will increase their canopy openness due to the removal of the largest trees. Forest fragmentation and climate warming will lead to a decrease in minimum humidity and an increase in maximum temperature. Fragmentation and warming are also likely to enhance each other's effects. In light of our results, all these factors are likely to contribute to a decrease in the abundance of epiphytes and thus lower the efficiency of the cloud forests as providers of clean water and maintainers of biodiversity.

4 Conclusions

The base cation gradient is relevant to the distribution of almost all of the studied Amazonian plant species. The lack of cation generalist species is remarkable. The cation niches of common species were not wider than those of rare species, which suggests that the niche breadth hypothesis may not apply to Amazonian ferns or Melastomataceae. The availability of habitat is likely to be an important driver of their distribution and commonness.

Adiantum, *Lindsaea* and Melastomataceae contain promising indicator species for the soil base cation concentration. In both groups, many species have niches that span less than 25 percent of the cation gradient, and each part of the gradient has many such species. Bioindication enables rapid soil mapping of the Amazonian rainforests. Since large parts of Amazonia are data-poor and difficult to access, and some are also urgently threatened by land-use changes, such rapid mapping techniques may be particularly useful for science-based land-use planning, beneficial to both agricultural production and nature conservation.

Forest structural differences between edaphically and floristically distinct Amazonian *terra firme* rainforests are not easy to measure on the ground, but new evidence from airborne measurements strongly suggests that they do exist. The Pebas Formation hosts forests with a more variable canopy than Nauta Formation forests. Structural differences likely contribute to differences in species composition, particularly that of understorey species and epiphytes.

Forest structure is an important driver of epiphyte abundance in cloud forests, partly because it affects microclimate. The responses of different epiphytic plant groups to changes in forest structure and microclimate are variable, but generally, logging and future climatic changes are likely to threaten the living conditions of epiphytes, which will have a negative impact on the water cycle in cloud forests.

Acknowledgements

This thesis project lasted more than a decade. Certainly, I did not plan it that way. My sincere thanks go to my supervisors, Hanna Tuomisto and Kalle Ruokolainen. I will perhaps never get to know what made you walk all this way with me. Your optimism and resilience are beyond those of many. Besides being excellent scientists, you are also invaluable friends.

I thank all members of the Amazon research team with whom I shared collaboration, walks in nature, joys and sometimes sorrows too. Glenda Cárdenas helped me a lot in organising my first field trip to Peru. With her help, I got acquainted with many good people in Iquitos and its surroundings. Johanna Toivonen and Carlos Gonzales Inca travelled with me twice to the Peruvian valley of Alto Mayo. Ulla Helimo was the person who eventually made the whole Alto Mayo cloud forest study possible, working in the coordination of the BioCuencas project. With Johanna, I shared exciting fieldwork, and also consultations about the renovation of old Finnish houses. Gabriela Zuquim taught me many things about Amazonia, Brazil, biology and life, and hosted me in Manaus. Mirkka Jones was always both an inspiration and a reliable source of R codes and solid scientific views. Calling Dr. Jones was a way to get through or around obstacles. Nelly Llerena and her family hosted me twice during my stay in Lima. Johanna Dávila, Gabriel Trujillo, Elias Mendivil, and several people from the Río Tahuayo and Alto Mayo areas provided invaluable fieldwork assistance. Anders Sirén gave many important insights about remote sensing, about life in general, and also shared his data.

I thank doctor Alberto Vicentini from INPA in Brazil and professor Jan Zelený from Taiwan University for pre-examining the thesis manuscript and giving constructive comments.

This thesis was rarely my only job. When it was, it was due to grants from Turun Yliopistosäätiö and the BioCuencas project, and I want to express my gratitude for that. Support from bosses at my primary job—high school teaching —was no less important. Since the thesis project was long, I had time to work in two different towns and different schools as well. I thank Timo Kouki, Kai Heino, and Juha-Markus Koistinen for letting me sometimes go away to work on this study.

I thank my long-time friends for discussions and company. Petteri Vihervaara took the time to read through the thesis manuscript at its final stage and give constructive comments. Together with Antti Rinne and Kimmo Lehtinen, we have shared good moments, been teachers, and reflected on the complexities of family life. You are precious fellows! With Ilkka Valve, we once studied through several classes at school and have remained friends since then.

During this project, my family grew to include three children. Doing PhD studies as a side-job would never have been possible without the help of the children's grandparents. My sincerest thanks go to Maija and Markku Niemi, and my parents, Marianne and Jaakko Suominen, for helping us along when there was just too much work to do.

Finally, Mai, I can write to you as you once wrote to me: You walked this way with me. Together with Kaisla, Kare and Lauha, we are a whole family. I thank you all, my dearest, for bearing with my frequent absence. We still have many happy days ahead.

References

Abraão, M.B., Nelson, B.W., Baniwa, J.C., Yu, D.W. and Shepard Jr, G.H. (2008). Ethnobotanical ground-truthing: indigenous knowledge, floristic inventories and satellite imagery in the upper Rio Negro, Brazil. *Journal of Biogeography*, 35, 2237–2248.

Adeney, J.M., Christensen, N.L., Vicentini, A. & Cohn-Haft, M. (2016). White-sand Ecosystems in Amazonia. *Biotropica*, 48, 7–23.

Ahmad, Z., Mulk Khan, S., Page, S., Alamri, S. & Hashem, M. (2022). Plants predict the mineral mines – A methodological approach to use indicator plant species for the discovery of mining sites. *Journal of Advanced Research*, 39, 119–133.

Anderson, A.B. (1981). White-Sand Vegetation of Brazilian Amazonia. *Biotropica*, 13 (3), 199–210.

Arellano, G., Cala, V. & Macía, M.J. (2014). Niche breadth of oligarchic species in Amazonian and Andean rain forests. *Journal of Vegetation Science*, 25 (6), 1355–1366.

Asner, G.P., Kellner, J.R., Kennedy-Bowdoin, T., Knapp, D.E., Anderson, C. & Martin, R.E. (2013). Forest canopy gap distributions in the Southern Peruvian Amazon. *PLoS One*, 8: e60875.

Assis, R.L., Haugaasen, T., Schöngart, J., Montero, J.C., Piedade, M.T.F. and Wittmann, F. (2015). Patterns of tree diversity and composition in Amazonian floodplain paleo-várzea forest. *Journal of Vegetation Science*, 26, 312–322.

Austin, M.P. (1976). On Non-Linear Species Response Models in Ordination. *Vegetatio* 33(1), 33–41.

Austin, M.P., Cunningham, R.B. & Fleming, P.M. (1984). New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. *Vegetatio* 55, 11–27.

Austin, M.P. & Smith, T.M. (1989). A new model for the continuum concept. *Vegetatio* 83(1), 35–47.

Austin, M.P., Nicholls, A.O. & Margules, C.R. (1990). Measurement of the realized qualitative niche - environmental niches of 5 Eucalyptus species. *Ecological Monographs*, 60 (2), 161–177.

Baraloto, C., Rabaud, S., Molto, Q., Blanc, L., Fortunel, C., Herault, B. et al. (2011). Disentangling stand and environmental correlates of aboveground biomass in Amazonian forests. *Global Change Biology*, 17, 2677-2688.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010). Regression analysis of spatial data. *Ecology Letters*, 13, 246–264.

Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C., ter Braak, C.J.F.T. (1990). Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 327 (1240), 263–278.

ter Braak, C.J.F. (1987) Weighted averaging using indicator values. *Data Analysis in Community and Landscape Ecology* (eds R.H. Jongman, O.F.R. van Tongeren & C.J.F. ter Braak), pp. 299. Pudoc, Wageningen.

ter Braak, C.J.F. & van Dam, H. (1989). Inferring pH from diatoms: a comparison of old and new calibration methods. *Hydrobiologia*, 178 (3), 209–223.

ter Braak, C.J.F. & Juggins, S. (1993). Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia, 269–270, 485–502.

Brown, J.H. (1984). On the Relationship between Abundance and Distribution of Species. *The American Naturalist*, 124 (2), 255–279.

Brown, N., Jennings, S., Wheeler, P., Nabe-Nielsen, J. (2000). An improved method for the rapid assessment of forest understorey light environments. *Journal of Applied Ecology*, 37: 1044-1053.

Bruijnzeel, L.A., Scatena, F.N. & Hamilton, L.S. (eds.) (2011). *Tropical montane cloud forests: science for conservation and management.* p. 5. Cambridge University Press, Cambridge, UK.

Bruno, J.F., Stachowicz, J.J. & Bertness, M.D. (2003). Inclusion of facilitation into ecological theory. *Trends in Ecology and Evolution*, 18(3), 119–125.

Caiafa, A.N. & Martins, F.R. (2010). Forms of rarity in tree species in the southern Brazilian Atlantic rainforest. *Biodiversity and Conservation,* 19, 2597–2618.

Cajander, A.K. (1926). The theory of forest types. *Acta Forestalia Fennica*, 29, 1–108.

Cámara-Leret, R., Tuomisto, H., Ruokolainen, K., Balslev, H., Munch Kristiansen, S. (2017). Modelling responses of western Amazonian palms to soil nutrients. *Journal of Ecology*, 105 (2), 367–381.

Cárdenas, G.G., Halme, K.J. & Tuomisto, H. (2007). Riqueza y Distribución Ecológica de Especies de Pteridofitas en la Zona del Río Yavarí-Mirín, Amazonía Peruana. *Biotropica*, 39 (5), 637–646.

Cárdenas, G.G., Tuomisto, H. & Lehtonen, S. (2016). Newly discovered diversity in the tropical fern genus *Metaxya* based on morphology and molecular phylogenetic analyses. *Kew Bulletin*, 71 (1), 1–27.

Choe, H., Thorne, J.H., Hijmans, R. & Seo, C. (2019). Integrating the Rabinowitz rarity framework with a National Plant Inventory in South Korea. *Ecology and Evolution*, 9, 1353–1363.

Chust, G., Chave, J., Condit, R., Aguilar, S., Lao, S. & Pérez, R. (2006). Determinants and spatial modeling of tree β-diversity in a tropical forest landscape in Panama. *Journal of Vegetation Science*, 17, 83–92.

Condit, R. Engelbrecht, Bettina M.J., Pino, D., Pérez, R. & Turner, B.L. (2013). Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *Proceedings of the National Academy of Sciences*, 110 (13), 5064–5068.

Costa, F.R.C. (2004). Structure and composition of the ground-herb community in a terra-firme Central Amazonian forest. *Acta Amazonica*, 34(1), 53–59.

Damasco, G., Vicentini, A., Castilho, C.V., Pimentel, T.P. & Nascimento, H.E.M. (2013). Disentangling the role of edaphic variability, flooding regime and topography of Amazonian white-sand vegetation. *Journal of Vegetation Science*, 24 (2), 384–394.

Damasco, G., Baraloto, C., Vicentini, A., Daly, D.C., Baldwin, B.G. & Fine, P.V.A. (2021). Revisiting the hyperdominance of Neotropical tree species under a taxonomic, functional and evolutionary perspective. *Scientific Reports*, 11 (1), 9585, 2021.

De Moura, Y.M., Hilker, T., Lyapustin, A.I., Galva, L.S., dos Santos, J.R., Anderson, L.O., de Sousa, C.H.R. & Arai, E. (2015). Seasonality and drought effects of Amazonian forests observed from multi-angle satellite data. *Remote Sensing of Environment*, 171, 278–290.

Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., Venail, P., Villéger, S. and Mouquet, N. (2010). Defining and measuring ecological specialisation. *Journal of Applied Ecology*, 47, 15–25.

DeWalt, S.J & Chave, J. (2004). Structure and biomass of four lowland Neotropical forests. *Biotropica*, 36, 7–19.

Diekmann, M. (2003). Species indicator values as an important tool in applied plant ecology - a review. *Basic and Applied Ecology*, 4, 493–506.

Doughty, C., Gaillard, C., Abraham, A., Burns, P., Keany, J., Aguirre-Gutierrez, J., Malhi, Y., Jantz, P., Koch, G., Shenkin, A. & Tang, H. (2023). Unstratified forests dominate the tropics especially in regions with lower fertility or higher temperatures. Preprint. *EcoEvoRxiv*. DOI: https://doi.org/10.32942/X2VC7T

Draper, F.C., Costa, F.R.C., Arellano, G. *et al.* (2021). Amazon tree dominance across forest strata. *Nature Ecology and Evolution*, 5, 757–767.

Duque, A., Sánchez, M., Cavelier, J. & Duivenvoorden, J.F. (2002). Different floristic patterns of woody understorey and canopy plants in Colombian Amazonia. *Journal of Tropical Ecology*, 18, 499–525.

Emilio, T., Nelson, B.W., Schietti, J., Desmoulière, S.J.-.-M., Espírito Santo, H.M.V. & Costa, F.R.C. (2010). Assessing the relationship between forest types and canopy tree beta diversity in Amazonia. *Ecography*, 33, 738–747.

Erwin, T.L. (1982). Tropical forests: Their richness in Coleoptera and other arthropod species. *The Coleopterists Bulletin*, 36 (1), 74–75.

Esquivel-Muelbert, A., Baker, T.R., Dexter, K.G., *et al.* (2017). Seasonal drought limits tree species across the Neotropics. *Ecography*, 40, 618–629.

Feldpausch, T., Banin, L., Phillips, O., Baker, T., Lewis, S., Quesada, C., Affum-Baffoe, K., Arets, E., Berry, N, Bird, M., Brondízio, E., Camargo, P., Chave, J., Djagbletey, G., Domingues, T., Drescher, M., Fearnside, P., França, M., Fyllas, N., Lloyd, . (2011). Height-diameter allometry of tropical forest trees. *Biogeosciences* 8, 1081–1106.

Fine, P.A., Daly, D.C. & Cameron, K.M. (2005), The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution*, 59, 1464–1478.

Fine, P.V.A., García-Villacorta, R., Pitman, N.C.A., Mesones, I. & Kembel, S.W. (2010). A Floristic Study of the White-Sand Forests of Peru. *Annals of the Missouri Botanical Garden,* 97(3), 283–305.

Frahm, J.-P. (1990). The effect of light and temperature on the growth of the bryophytes of tropical rain forests. *Nova Hedwigia*, 51, 151–164.

Frahm, J.-P. & Gradstein, S.R. (1991). An altitudinal zonation of tropical rain forests using bryophytes. *Journal of Biogeography*, 18 (6), 669–678.

Gaston, K.J. & Lawton, J.H. (1990). The population ecology of rare species. Journal of Fish Biology, 37(Suppl.A), 97–104.

Gauch, H.G. & Whittaker, R.H. (1972). Coenocline simulation. *Ecology*, 53 (3), 446–451.

Gégout, J.-C., Hervé, J.-C., Houllier, F. & Pierrat, J.-C. (2003). Prediction of forest soil nutrient status using vegetation. *Journal of Vegetation Science*, 1,: 55–62.

Gehrig-Downie, C, Obregón, A., Bendix, J. & Gradstein, S.R. (2011). Epiphyte Biomass and Canopy Microclimate in the Tropical Lowland Cloud Forest of French Guiana. *Biotropica*, 43 (5), 591–596.

Gentry, A.H. (1982) Patterns of neotropical plant-species diversity. *Evolutionary Biology*, 15, 1–85.

Gentry, A. H. (1986). Endemism in tropical vs temperate plant communities. In Conservation Biology: The Science of Scarcity and Diversity, ed. Soulé, M.. Sunderland, MA. USA: Sinauer Associates, pp. 153–81.

Gentry, A.H. & Dodson, C.H. (1987). Diversity and biogeography of neotropical epiphytes. *Annals of the Missouri Botanical Garden*, 74, 205–233.

Gradstein, R.S. (2008). Epiphytes of tropical montane forests – impact of deforestation and climate change. In: Gradstein, R.S., Homeier, J. & Gansert, D. (eds.) *Biodiversity and ecology series 2. The tropical mountain Forest – Patterns and processes in a biodiversity hotspot.* Göttingen Centre for Biodiversity and Ecology, Göttingen, Germany.

Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006). Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters,* 9, 399–409.

Guevara Andino, J.E., Pitman, N.C.A., ter Steege, H. *et al.* (2021). The contribution of environmental and dispersal filters on phylogenetic and taxonomic beta diversity patterns in Amazonian tree communities. *Oecologia*, 196, 1119–1137.

Guèze, M., Paneque-Gálvez, J., Luz, A.C., Pino, J., Orta-Martínez, M., Reyes-García, V. & Macía, M.J. (2013). Determninants of tree species turnover in a southern Amazonian rain forest. *Journal of Vegetation Science*, 24, 284–295.

Hájková, P., Hájek, M., Apostolova, I., Zelený, D. & Dítě, D. (2008). Shifts in the ecological behaviour of plant species between two distant regions: evidence from the base richness gradient in mires. *Journal of Biogeography*, 35, 282–294.

Halme, K. & Bodmer, R. Correspondence between Scientific and Traditional Ecological Knowledge: Rain Forest Classification by the Non-Indigenous Ribereños in Peruvian Amazonia. *Biodiversity and Conservation*, 16, 1785–1801.

Hanski, I. (1982). Dynamics of Regional Distribution: the Core and Satellite Species Hypothesis. *Oikos*, 38(2), 210.

Harpole, W.S. & Tilman, D. (2006). Non-neutral patterns of species abundance in grassland communities. *Ecology Letters*, 9, 15–23.

Heikkinen, J. & Mäkipää, R. (2010). Testing hypotheses on shape and distribution of ecological response curves. *Ecological Modelling*, 221(3), 388–399.

Hedwall, P.-O., Brunet, J. & Diekmann, M. (2019). With Ellenberg indicator values towards the north: Does the indicative power decrease with distance from Central Europe? *Journal of Biogeography*, 46, 1041–1053.

Hengl, T. et al. 2017. SoilGrids250m: global gridded soil information based on machine learning. – PLoS One 12: 1–40.

Higgins, M.A. & Ruokolainen, K. (2004). Rapid tropical forest inventory: A comparison of techniques based on inventory data from western Amazonia. *Conservation Biology*, 18(3), 799–811.

Higgins, M.A., Ruokolainen, K., Tuomisto, H., Llerena, N., Cardenas, G., Phillips, O.L., Vásquez, R. & Räsänen, M. (2011). Geological control of floristic composition in Amazonian forests. *Journal of Biogeography*, 38, 2136–2149.

Higgins, M.A., Asner, G.P., Perez, E., Elespuru, N., Tuomisto, H., Ruokolainen, K. & Alonso, A. (2012). Use of Landsat SRTM Data to detect broad-scale biodiversity patterns in Northwestern Amazonia. *Remote Sensing*, 4, 2401–2418.

Higgins, M.A., Asner, G.P., Anderson, C.B., Martin, R.E., Knapp, D.E., Tupayachi, R., Perez, E., Elespuru, N. & Alonso, A. (2015). Regional Scale Drivers of Forest Structure and Function in Northwestern Amazonia. *PLoS ONE*, 10(3): e0119887.

Holz, I. & Gradstein, S.R. (2005). Cryptogamic epiphytes in primary and recovering upper montane oak forests of Costa Rica - species richness, community composition and ecology. *Plant Ecology,* 178, 89–109.

Honorio Coronado, E. N., Baker, T. R., Phillips, O. L., Pitman, N. C. A., Pennington, R. T., Vásquez Martínez, R., Monteagudo, A., Mogollón, H., Dávila Cardozo, N., Ríos, M., García-Villacorta, R., Valderrama, E., Ahuite, M., Huamantupa, I., Neill, D. A., Laurance, W. F., Nascimento, H. E. M., Soares de Almeida, S., Killeen, T. J., Arroyo, L., Núñez, P., & Freitas Alvarado, L. (2009). Multi-scale comparisons of tree composition in Amazonian terra firme forests. *Biogeosciences*, 6, 2719–2731.

Hoorn, C. (1993). Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia - results of a palynostratigraphic study. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 105, 267–309.

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

Huisman J., Olff H., Fresco L.F.M. (1993). A Hierarchical Set of Models for Species Response Analysis. *Journal of Vegetation Science,* 4, 37–46.

Hutchinson, G.E. (1957). The multivariate niche. *Cold Spr. Harb. Symp. Quant. Biol.* 22, 415–421.

IBGE (2004). Mapa de vegetação do Brasil. 3rd edn. Fundação Instituto Brasileiro de Geografia e Estatística. Ministério da Agricultura, Rio de Janeiro.

Jabot, F., Etienne, R.S. & Chave, J. (2008). Reconciling neutral community models and environmental filtering: theory and an empirical test. *Oikos*, 117, 1308–1320.

Jansen, F. & Oksanen, J. (2013). How to model species responses along ecological gradients – Huisman–Olff–Fresco models revisited. *Journal of Vegetation Science*, 24 (6), 1108-1117.

John, R., Dalling, J.W., Harms, K.E., Yavitt, J.B., Stallard, R.F., Mirabello, M., Hubbell, S.P., Valencia, R., Navarrete, H., Vallejo, M. & Foster, R.B. (2007). Soil nutrients influence spatial distributions of tropical tree species. *PNAS*, 104(3), 864–869.

Jones, M.M., Tuomisto, H. & Olivas, P.C. (2008). Differences in the degree of environmental control on large and small tropical plants: just a sampling effect? *Journal of Ecology*, 96 (2), 367–377.

Kalliola, R., Puhakka, M., Salo J., Tuomisto, H. & Ruokolainen, K. (1991). The Dynamics, Distribution and Classification of Swamp Vegetation in Peruvian Amazonia. *Annales Botanici Fennici*, 28 (3), 225–239.

Karger, D.N., Kluge, J., Abrahamczyk, S., Salazar, L., Homeier, J., Lehnert, M., Amoroso, V.B. & Kessler, M. (2012). Bryophyte cover on trees as proxy for air humidity in the tropics. *Ecological Indicators*, 20, 277–281.

Katila, M. & Tomppo, E. (2001). Selecting estimation parameters for the Finnish multisource National Forest Inventory. *Remote Sensing of Environment*, 76, 16–32.

Kinupp, V.F. & Magnusson, W.E. (2005). Spatial patterns in the understorey shrub genus *Psychotria* in central Amazonia: effects of distance and topography. *Journal of Tropical Ecology*, 21 (4), 363–374.

Kissling, W.D. & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, 17, 59–71.

Kraft, N.J.B. & Ackerly, D.D. (2014). Assembly of plant communities. In: *Ecology and the Environment*, pp. 67–88, Springer, New York.

Kristiansen, T., Svenning, J-C., Eiserhardt, W.L., Pedersen, D., Brix, H., Munch Kristiansen, S., Knadel, M., Grández, C. & Balslev, H. (2012). Environment versus dispersal in the assembly of western Amazonian palm communities. *Journal of Biogeography*, 39 (7), 1318–1322.

Krömer T. & Gradstein, S.R. (2004). Species richness of vascular epiphytes in two primary forests and fallows in the Bolivian Andes. *Selbyana,* 24, 190–195.

Krömer, T., Kessler, M. & Gradstein, S.R. (2007). Vertical stratification of vascular epiphytes in submontane and montane forest of the Bolivian Andes: the importance of the understory. *Plant Ecology*, 189, 261–178.

Kuusipalo, J. (1985). An ecological study of upland forest site classification in southern Finland. *Acta Forestalia Fennica*, 192, 1–77.

Lawesson, J.E. & Oksanen, J. (2002). Niche characteristics of Danish woody species as derived from coenoclines. *Journal of Vegetation Science*, 13, 279–290.

Lehtonen, S., Jones, M.M., Zuquim, G., Prado, J. & Tuomisto, H. (2015). Phylogenetic relatedness within Neotropical fern communities increases with soil fertility. *Global Ecology and Biogeography*, 24 (6), 695–705.

Lehtonen, S., Muscarella, R., Moulatlet, G., Balslev, H. & Tuomisto, H. (2021). Edaphic heterogeneity and the evolutionary trajectory of Amazonian plant communities. *Ecology and Evolution*, 11, 17672–17685.

León Vargas, Y., Engwald, S. & Proctor, M.C.F. (2006). Microclimate, light adaptation and desiccation tolerance of epiphytic bryophytes in two Venezuelan cloud forests. *Journal of Biogeography*, 33, 901–913.

Levis C., Souza, P.Fd., Schietti, J., Emilio, T., Pinto J.L.P.dV., Clement, C.R., et al. (2012). Historical Human Footprint on Modern Tree Species Composition in the Purus-Madeira Interfluve, Central Amazonia. *PLoS ONE*, 7(11), e48559.

Lovejoy, T.E. & Nobre, C. (2019). Amazon tipping point: Last chance for action. *Science Advances*, 5 (12).

Lu, D., Mausel, P., Brondízio, E. & Moran, E. (2004). Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. *Forest Ecology and Management*, 198, 149–167.

Lähteenoja, O., and Page, S. (2011). High diversity of tropical peatland ecosystem types in the Pastaza-Marañón basin, Peruvian Amazonia, *Journal of Geophysical Research*, 116, G02025.

McMichael, C.H., Palace, M.W. and Golightly, M. (2014). Bamboo-dominated forests and pre-Columbian earthwork formations in south-western Amazonia. *Journal of Biogeography*, 41, 1733–1745.

McNaughton, S.J. & Wolf, L.L. (1970). Dominance and the niche in ecological systems. *Science*, 167(3915), 131–139.

MacArthur, R.H. & MacArthur, J.W. (1961). On Bird Species Diversity. *Ecology*, 42 (3), 594–598.

Males, J. (2016). Think tank: water relations of Bromeliaceae in their evolutionary context. *Botanical Journal of the Linnean Society*, 181, 415–440.

Malhi, Y., Wood, D., Baker, T.R., Wright, J., Phillips, O.L., Cochrane, T., Meir, P., Chave, J., Almeida, S., Arroyo, L., Higuchi, N., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Monteagudo, A., Neill, D.A., Vargas, P.N., Pitman, N.C.A., Quesada, C.A., Salomao, R., Silva, J.N.M., Lezama, A.T., Terborgh, J., Martinez, R.V. & Vinceti, B. (2006). The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology*, 12, 1107–1138.

Mantovani, A. (1999). Leaf morpho-physiology and distribution of epiphytic aroids along a vertical gradient in a Brazilian rainforest. *Selbyana*, 20 (2), 241–249.

Marca-Zevallos, M.J., Moulatlet, G.M., Sousa, T.R., *et al.* (2022), Local hydrological conditions influence tree diversity and composition across the Amazon basin. *Ecography*, 2022: e06125.

Martin, C.E. & Adamson, V.J. 2001. Photosynthetic capacity of mosses relative to vascular plants. *Journal of Bryology*, 23, 319–323.

Michaelis, J. & Diekmann, M.R. (2018). Biased niches – Species response curves and niche attributes from Huisman-Olff-Fresco models change with differing species prevalence and frequency. *PLOS ONE,* 12(8), e0183152.

Montgomery, R. & Chazdon, R. (2002). Light gradient partitioning by tropical tree seedlings in the absence of canopy gaps. *Oecologia*, 131(2), 165–174.

Moulatlet, G.M., Costa, F.R.C., Rennó, C.D., Emilio, T. & Schietti, J. (2014). Local Hydrological Conditions Explain Floristic Composition in Lowland Amazonian Forests. *Biotropica* 46 (4), 395–403.

Nadkarni, N.M. (1984). Epiphytic biomass and nutrient capital in a Neotropical elfin forest. *Biotropica*, 16 (4), 249–256.

Nadkarni, N.M. (2000). Colonization of stripped branch surfaces by epiphytes in a lower montane cloud forest, Monteverde, Costa Rica. *Biotropica*, 32, 358–363.

Nadkarni, N.M. & Solano, R. (2002). Potential effects of climate change on canopy communities in a tropical cloud forest: an experimental approach. *Ecologia,* 131, 580–594.

Nicotra, A.B., Chazdon, R.L. & Iriarte, S.V.B. (1999). Spatial heterogeneity of light and woody seedling regeneration in tropical wet forests. *Ecology*, 80, 1908–1926.

Nöske, N.N., Hilt, N., Werner, F.A., Brehm, G., Fiedler, K., Sipman, H.J.M. & Gradstein, S.R. (2008). Disturbance effects on diversity of epiphytes and moths in a montane forest in Ecuador. *Basic and Applied Ecology*, 9, 4–12.

Oksanen, J. & Minchin, P. (2002). Continuum theory revisited: what shapes are species responses on ecological gradients? *Ecological Modelling*, 157(2–3), 119–129.

Ozanne, C.M.P., Anhuf, D., Boulter, S.L., Keller, M., Kitching, R.L., Körner, C., Meinzer, F.C., Mitchell, A.W., Nakashizuka, T., Silva Dias, P.L., Stork, N.E., Wright, S.J. & Yoshimura, M. (2003). Biodiversity Meets the Atmosphere: A Global View of Forest Canopies. *Science,* 301, 183–186.

Paoli, G.D., Curran, L.M. & Zak, D.R. (2006). Soil nutrients and beta diversity in the Bornean Dipterocarpaceae: evidence for niche partitioning by tropical rain forest trees. *Journal of Ecology*, 94 (1), 157–170.

Paoli, G.D., Curran, L.M., & Slik, J.W.F. (2008). Soil nutrients affect spatial patterns of aboveground biomass and emergent tree density in southwestern Borneo. *Oecologia*, 155, 287–299.

Pearson, S. (1988). Botanical Indicators in Historical Archaeology. *Australian Journal of Historical Archaeology*, 6, 74–82.

Petter, G., Wagner, K, Wanek, W., Sanchez Delgado, E.J., Zotz, G., Sarmento Cabral, J. & Kreft, H. (2016). Functional leaf traits of vascular epiphytes: vertical trends within the forest, intra- and interspecific trait variability, and taxonomic signal. *Functional Ecology*, 30, 188–198.

Phillips, O.L., Vargas, P.N., Monteagudo, A.L., Cruz, A.P., Zans, M.-E.C., Sánchez, W.G., Yli-Halla, M. & Rose, S. (2003). Habitat association among Amazonian tree species: A landscape-scale approach. *Journal of Ecology* 91(5), 757–775.

Phillips, O. L., Baker, T. R., Arroyo, L., Higuchi, N., Killeen, T. J., Laurance, W. F., Lewis, S. L., Lloyd, J., Malhi, Y., Monteagudo, A., Neill, D. A., Nüñez Vargas, P., Silva, J. N. M., Terborgh, J., Vásquez Martínez, R., Alexiades, M., Almeida, S., Brown, S., Chave, J., Comiskey, J. A., Czimczik, C. I., Di Fiore, A., Erwin, T., Kuebler, C., Laurance, S. G., Nascimento, H. E. M., Olivier, J., Palacios, W., Patiño, S., Pitman, N. C. A., Quesada, C. A., Saldias, M., Torres Lezama, A., & Vinceti, B. (2004). Pattern and process in Amazon tree turnover, 1976–2001. *Philos. T. R. Soc. B*, 359, 381–407.

Pitman, N.C.A., Terborgh, J., Silman, M.R. & Nuñez, V.P. (1999). Tree species distributions in an upper Amazonian forest. *Ecology,* 80 (8), 2651–2661.

Pitman, N.C.A., Terborgh, J.W., Silman, M.R., Núñez, V.P., Neill, D.A., Cerón, C.E., Palacios, W.A. & Aulestia, M. (2001). Dominance and Distribution of Tree Species in Upper Amazonian Terra Firme Forests. *Ecology*, 82 (8), 2101–2117.

Pitman, N.C.A., Silman, M.R. & Terborgh, J.W. (2013). Oligarchies in Amazonian tree communities: a ten-year review. *Ecography*, 36 (2), 114–123.

Pomara, L.Y., Ruokolainen, K., Tuomisto, H. & Young, K.R. (2012). Avian Composition Co-varies with Floristic Composition and Soil Nutrient Concentration in Amazonian Upland Forests. *Biotropica*, 44 (4), 545–553.

Quesada, C. A., Phillips, O. L., Schwarz, M., Czimczik, C. I., Baker, T. R., Patiño, S., ... & Lloyd, J. (2012). Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. Biogeosciences, 9(6), 2203-2246.

Rabinowitz, D. (1981). Seven forms of rarity. Pages 205–217 in H. Synge, editor. The biological aspects of rare plant conservation. Wiley, New York, New York, USA.

RADAMBRASIL (1978). Projeto RADAMBRASIL. Vol. (1:34). Geologia, geomorfologia, pedologia, vegetação e uso potencial da terra. Departamento Nacional de Produção Mineral. Rio de Janeiro, BR.

Rebata, L.A., Gingras, M.K., Räsänen, M.E. & Barberi, M. (2006). Tidal-channel deposits on a delta plain from the Upper Miocene Nauta Formation, Marañón Foreland Sub-basin, Peru. *Sedimentology*, 53, 971–1013.

Reinecke, J., Wulf, M., Baeten, L., Brunet, J., Decocq, G., De Frenne, P., Diekmann, M., Graae, B.J., Heinken, T., Hermy, M., Jamoneau, A., Lenoir, J., Plue, J., Orczewska, A., Van Calster, H., Verheyen, K. & Naaf, T. (2016). Acido- and neutrophilic temperate forest plants display distinct niche shifts in ecological pH niche across north-western Europe. *Ecography,* 39, 1164–1175.

Rosindell, J., Hubbell, S.P. & Etienne, R.S. (2011). *The Unified Neutral Theory of Biodiversity and Biogeography* at Age Ten. *Trends in Ecology and Evolution*, 26 (7), 340–348.

Ruokolainen, K., Linna, A. & Tuomisto, H. (1997). Use of Melastomataceae and Pteridophytes for Revealing Phytogeographical Patterns in Amazonian Rain Forests. *Journal of Tropical Ecology*, 13, 243–256.

Ruokolainen, K. & Tuomisto, H. (1998). Vegetación natural de la zona de Iquitos. In: Kalliola, R.; Flores Paitán, S. (Ed.). Geoecología y desarrollo amazónico: estudio integrado en la zona de Iquitos, Perú. University of Turku, Finland. p. 253–365.

Ruokolainen, K. & Vormisto, J. (2000). The most widespread Amazonian palms tend to be tall and habitat generalists. *Basic and Applied Ecology*, 1 (2), 97–108.

Ruokolainen, K., Tuomisto, H., Macía, M.J., Higgins, M.A. & Yli-Halla, M. (2007). Are floristic and edaphic patterns in Amazonian rain forests congruent for trees, pteridophytes and Melastomataceae? *Journal of Tropical Ecology*, 23, 13-25.

Russo, S.E., Zhang, L., Tan, S. (2012). Covariation between understorey light environments and soil resources in Bornean mixed dipterocarp rain forest. *Journal of Tropical Ecology*, 28, 33-44.

Rydgren, K., Økland, R.H. and Økland, T. (2003). Species response curves along environmental gradients. A case study from SE Norwegian swamp forests. *Journal of Vegetation Science*, 14, 869–880.

Salo, J., Kalliola, R., Häkkinen, I., Mäkinen, Y., Niemelä, P., Puhakka, M. & Coley, P.D. (1986). River dynamics and the diversity of Amazon lowland forest. *Nature*, 322 (6076), 254–258.

Salovaara, K.J., Cárdenas, G.G. & Tuomisto, H. (2004). Forest classification in an Amazonian rainforest landscape using pteridophytes as indicator species. *Ecography*, 27, 689–700.

Salovaara, K., Thessler, S., Malik, R.N. & Tuomisto, H. (2005). Classification of Amazonian primary rainforest vegetation using Landsat ETM+ satellite imagery. *Remote Sensing of Environment*, 97 (1), 39–51.

Sanchez, P.A. and Buol, S.W. (1974). Properties of Some Soils of the Upper Amazon Basin of Peru. Soil Science Society of America Journal, 38, 117–121.

Scheffers, B.R., Phillips, B.L., Laurance, W.F., Sodhi, N.S., Diesmos, A. & Williams, S.E. (2013). Increasing arboreality with altitude: a novel biogeographic dimension. *Proceedings of the Royal Society B,* 280, 20131581.

Schmidt, G., Stuntz, S. & Zotz, G. (2001). Plant size: an ignored parameter in epiphyte ecophysiology? *Plant Ecology*, 153, 65–72.

Shepard Jr., G.H., Yu, D.W., Lizarralde, M. & Italiano, M. (2001). Rain forest habitat classification among the Matsigenka of the Peruvian Amazon. *Journal of Ethnobiology*, 21(1), 1–38.

Silvertown, J. (2004). Plant coexistence and the niche. *Trends in Ecology and Evolution*, 19(11), 605–611.

Sirén, A., Tuomisto, H. & Navarrete, H. (2013). Mapping environmental variation in lowland Amazonian rainforests using remote sensing and floristic data. *International Journal of Remote Sensing*, 34 (5), 1561–1575.

Slatyer, R.A., Hirst, M. & Sexton, J.P. (2013). Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters,* 16, 1104–1114.

Slik, J.W.F., Aiba, S.I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K. *et al*. (2010). Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. *Global Ecology and Biogeography*, 19, 50–60.

Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10, 1115–1123.

ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.-F., Prévost, M.-F., Spichiger, R., Castellanos. H., von Hildebrand, P. & Vásquez, R. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, 443 (7110), 444–447.

ter Steege, H. *et al.* (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342 (6156), 1243092.

ter Steege, H., Prado, P.I., Lima, R.A.F.d. *et al.* (2020). Biased-corrected richness estimates for the Amazonian tree flora. *Scientific Reports*, 10, 10130 (2020).

Symonds, M.R.E. & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioral ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65, 13–21.

Thornthwaite, C. (1948). An approach toward a rational classification of climate. *Geographical Review*, 38 (1), 55–94.

Tobón, C., Köhler, L., Frumau, K.F.A., Bruijnzeel, L.A., Burkard, R. & Schmid, S. (2010). Water dynamics of epiphytic vegetation in a lower montane cloud forest: fog interception, storage, and evaporation. In: Bruijnzeel, L.A., Scatena, F.N. & Hamilton, L.S. (eds.) *Tropical montane cloud* *forests: science for conservation and management,* pp. 261–267. Cambridge University Press, Cambridge, UK.

Toledo, M., Peña-Claros, M., Bongers, F., Alarcón, A., Balcázar, J., Chuviña, J., Leaño, C., Licona, J.C. & Poorter, L. (2012). Distribution patterns of tropical woody species in response to climatic and edaphic gradients. *Journal of Ecology*, 100 (1), 253–263.

Tuomisto, H. & Ruokolainen, K. (1994). Distribution of Pteridophyta and Melastomataceae along an edaphic gradient in an Amazonian rain forest. *Journal of Vegetation Science*, 5 (1), 25–34.

Tuomisto, H., Ruokolainen, K., Kalliola, R., Linna, A., Danjoy, W. & Rodriguez, Z. (1995). Dissecting Amazonian biodiversity. *Science*, 269, 63–66.

Tuomisto, H. & Poulsen, A.D. (1996). Influence of edaphic specialization on pteridophyte distribution in neotropical rain forests. *Journal of Biogeography*, 23, 283–293.

Tuomisto, H., Ruokolainen, K., Poulsen, A.D., Moran, R.C., Quintana, C., Cañas, G. and Celi, J. (2002). Distribution and Diversity of Pteridophytes and Melastomataceae along Edaphic Gradients in Yasuní National Park, Ecuadorian Amazonia. *Biotropica*, 34, 516–533.

Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003 a). Dispersal, Environment, and Floristic Variation of Western Amazonian Forests. *Science*, 299 (5604), 241–244.

Tuomisto, H., Poulsen, A.D., Ruokolainen, K., Moran, R.C., Quintana, C., Celi, J. & Cañas, G. (2003 b). Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. *Ecological Applications*, 13 (2), 352–371.

Tuomisto, H., Ruokolainen, K., Aquilar, M. & Sarmiento, A. (2003 c). Floristic patterns along a 43-km long transect in an Amazonian rain forest. *Journal of Ecology*, 91, 743–756.

Tuomisto, H. (2006). Edaphic niche differentiation among Polybotrya ferns in western Amazonia: implications for coexistence and speciation. *Ecography*, 29, 273–284.

Tuomisto, H., Moulatlet, G.M, Balslev, H., Emilio. T., Figueiredo, F.O.G., Pedersen, D. & Ruokolainen, K. (2016). A compositional turnover zone of biogeographical magnitude within lowland Amazonia. *Journal of Biogeography*, 43 (12), 2400–2411.

Tuomisto, H., Van doninck, J., Ruokolainen, K., Moulatlet, G.M., Figueiredo, F.O.G., Sirén, A., Cárdenas, G., Lehtonen, S. & Zuquim, G. (2019). Discovering floristic and geoecological gradients across Amazonia. *Journal of Biogeography* 46, 1734–1748.

Veloso, H.P., Japiassu, A.M.S., Goes Filho, L. & Leite, P.F. (1974). As regiões fitoecológicas, sua natureza e seus recursos económicos. Departamento Nacional de Produção Mineral Río de Janeiro, Brazil. p. 1–119.

Veneklaas, E.J. & van Ek, R (1990). Rainfall interception in two tropical montane rain forests, Colombia. *Hydrological Processes*, 4, 311–326.

Vormisto, J., Tuomisto, H. & Oksanen, J. (2004). Palm distribution patterns in Amazonian rainforests: what is the role of topographic variation? *Journal of Vegetation Science*, 15, 485–494.

Webb, L. J., Tracey, J. G., Williams, W. T., & Lance, G. N. (1967). Studies in the Numerical Analysis of Complex Rain-Forest Communities: II. The Problem of Species-Sampling. *Journal of Ecology*, 55(2), 525–538.

Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30 (3), 279–338.

Wilson, S. McG., Pyatt, D.G., Malcolm, D.C. & Connolly, T. (2001). The use of ground vegetation and humus type as indicators of soil nutrient regime for an ecological site classification of British forests. *Forest Ecology and Management*, 140 (2–3).

Wolf, J.H.D. (1993). Diversity patterns and biomass of epiphytic bryophytes and lichens along an altitudinal gradient in the northern Andes. *Annals of the Missouri Botanical Garden*, 80, 928–960.

Wolf, P.G., Schneider, H. & Ranker, T.A. (2001). Geographic distributions of homosporous ferns: does dispersal obscure evidence of vicariance? *Journal of Biogeography*, 28 (2), 263–270.

Zotz, G. (2013). The systematic distribution of vascular epiphytes - a critical update. *Botanical Journal of the Linnean Society*, 171, 453–481.

Zotz, G. & Bader, M.Y. (2009). Epiphytic plants in a changing world: global change effects on vascular and nonvascular epiphytes. In: Lüttge, U., Beyschlag, W., Büdel, B. & Francis, D. (eds.) *Progress in botany*, vol. 70, pp. 147–170. Springer-Verlag, Berlin, DE.

Zotz, G., Bündel, B., Meyer, A., Zellner, H. & Lange, O.L. (1997). Water relations and CO2 exchange of tropical bryophytes in a lower montane rain forest in Panama. *Botanica Acta*, 110, 9–17.

Zuquim, G., Costa, F.R.C., Prado, J. & Braga-Neto, R. (2009). Distribution of pteridophyte communities along environmental gradients in Central Amazonia, Brazil. *Biodiversity and Conservation*, 18 (1), 151–166.

Zuquim, G., Tuomisto, H., Costa, F.R.C., Prado, J., Magnusson, W.E., Pimentel, T., Braga-Neto, R., Figueiredo, F.O.G. (2012). Broad scale distribution of ferns and lycophytes along environmental gradients in central and northern Amazonia, Brazil. *Biotropica,* 44 (6), 752-762.

Zuquim, G., Tuomisto, H., Jones, M.M., Prado, J., Figueiredo, F.O.G., Moulatlet, G.M., Costa, F.R.C., Quesada, C.A. & Emilio, T. (2014). Predicting environmental gradients with fern species composition in Brazilian Amazonia. *Journal of Vegetation Science*, 25 (5), 1195–1207.

Zwieniecki, M.A. & Boyce, C.K. (2014). Evolution of a unique anatomical precision in angiosperm leaf venation lifts constraints on vascular plant ecology. *Proceedings of the Royal Society B*, 281, 20132829.

TURUN YLIOPISTO
UNIVERSITY OF TURKU

Painosalama, Turku, Finland 2023

ISBN 978-951-29-9298-0 (PRINT)
ISBN 978-951-29-9299-7 (PDF)
ISSN 0082-6979 (Print)
ISSN 2343-3183 (Online)