Universiteit Leiden
The Netherlands

Borneo : a quantitative analysis of botanical richness, endemicity and floristic regions based on herbarium records

Raes, N.

Citation

Raes, N. (2009, February 11). *Borneo : a quantitative analysis of botanical richness, endemicity and floristic regions based on herbarium records*. Retrieved from https://hdl.handle.net/1887/13470

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/13470

Note: To cite this publication please use the final published version (if applicable).

Borneo

A quantitative analysis of botanical richness,
endemicity and floristic regions
based on herbarium records

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College van Promoties
te verdedigen op woensdag 11 februari 2009
klokke 15.00 uur

door

NIELS RAES
Geboren te Bladel
in 1971

Promotiecommisie:

Promotor:	Prof. dr. P. Baas
Copromotoren:	Dr. H. ter Steege Dr. M.C. Roos
Referent:	Prof. dr. F.J.J.M. Bongers (Wageningen University)
Overige leden:	Prof. dr. D.J. Mabberley Prof. dr. E.F. Smets Prof. dr. M.S.M. Sosef Prof. dr. P.C. van Welzen

A quantitative analysis of botanical richness, endemicity and floristic regions based on herbarium records

Niels Raes

NATIONAAL HERBARIUM NEDERLAND
Leiden University branch
2009

Raes, N.
Borneo: A quantitative analysis of botanical richness, endemicity and floristic regions based on herbarium records.
ISBN: 978-90-8570-318-1
Printing: Wöhrmann Print Service
Graphic Design: Ed van Oosterhout
Eindredactie Nederlandse samenvatting: Koen-Machiel van de Wetering
FSC Mixed Sources: This book is printed on FSC certified paper.

Cover: The symbol used on the cover and throughout the thesis is a Dayak symbol that stands for a man figure. The precise meaning of the symbol depends on the context in which it is placed, but generally it depicts the continuity of (human) life.

Chapter 1: Partly published in Blumea 54(1), N. Raes and P.C. van Welzen, The demarcation and internal division of Flora Malesiana: 1857 – Present, ©2009, with permission from the Nationaal Herbarium Nederland, Leiden University branch.
Chapter 2: Reprinted from Blumea 54(1), N. Raes, J.B. Mols, L.P.M. Willemse and E.F. Smets, Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images, ©2009, with permission from the Nationaal Herbarium Nederland, Leiden University branch.
Chapter 3: Reprinted from Ecography 30, N. Raes and H. ter Steege, A null-model for significance testing of presence-only species distribution models, ©2007, with permission from the Oikos Editorial Office.
Chapter 4: Reprinted from Ecography (accepted), N. Raes, M.C. Roos, J.W.F. Slik, E.E. van Loon and H. ter Steege, Botanical richness and endemicity patterns of Borneo derived from species distribution models, ©2009, with permission from the Oikos Editorial Office.
Chapter 5: Submitted to Journal of Biogeography
Chapter 6: In preparation
Remainder of the thesis ©2009, Nationaal Herbarium Nederland, Leiden University branch.

No part of this publication, apart from bibliographic data and brief annotations in critical reviews, may be reproduced, re-recorded or published in any form, including print, photocopy, microform, electronic or electromagnetic record without written permission by the publishers.

Summary

Although it is widely recognized that Borneo harbours one of the world's most important biodiversity hotspots, the spatial patterns of botanical richness, endemicity, 'centres of endemicity', and Borneo's floristic regions, until now have largely been based on informal expert opinion. Recent digitization of the botanical collections of Borneo, housed at the National Herbarium of the Netherlands, has provided a database that allowed a quantitative, spatial analysis of the components of botanical diversity of Borneo.

The objectives of this study are to develop high-resolution spatial maps of the patterns of botanical richness, -endemicity, 'centres of endemicity', and the floristic regions of Borneo derived from species distribution models. The resulting maps are related to environmental conditions to explain the patterns of the different components of botanical diversity and the recognized floristic regions. Finally, we assess the extent to which areas of high botanical diversity and the different floristic regions remain forested today to guide conservation efforts of the threatened forests of Borneo.

We used 'Maxent' to develop species distribution models for species treated in 'Flora Malesiana' and represented by at least five records. The 2273 species distribution models were statistically tested with a method we developed for this purpose, resulting in 1439 significant models (63.3%), covering 8577 grid cells (5 arc-minute resolution, ca. 100 km2) of Borneo.

The 1439 significant models were superimposed to generate patterns of botanical richness, -endemicity, and 'centres of endemicity'. The highest botanical richness is predicted to occur in northern and northwestern Borneo. The northern Crocker Mountains range with Mount Kinabalu, and the high mountains of central East Kalimantan have the highest botanical endemicity values. The 'centres of endemicity' are found on Mount Kinabalu and the northern Crocker Range Mountains, the southern Müller Mountains, the east side of the Meratus Mountains, and the Sangkulirang peninsula. Areas of high botanical richness and -endemicity are characterized by a relatively small range in annual temperature, but with seasonality within that range. Furthermore, these areas are least affected by the El Niño Southern Oscillation drought events. 'Centres of endemicity' are characterized by ecological distinctiveness in altitude, edaphic conditions, annual precipitation, or a combination of these factors.

The 11 floristic regions of Borneo were recognized based on a presence/absence matrix derived from the 1439 significant species distribution models. This matrix was analysed using a hierarchical cluster analysis, and the resulting cluster dendrogram was pruned using indicator species analysis (ISA) to partition floristic regions. This method allowed the quantitative confirmation of the floristic distinctiveness and extent of montane rain forest, kerangas, peat swamp, and fresh water swamp forest. The lowland rain forest, previously recognized as one floristic region was divided in at least four (and possibly six) different floristic regions, viz. the lowlands of (i) Sabah and Sarawak, (ii) East Kalimantan, (iii) southern Borneo, and (iv) the 'Wet hill forest of Sarawak'. We could not distinguish, but do recognize, the 'Kinabalu highlands', mangroves, and forests on limestone and ultramafic rock due to the spatial resolution (100 km2) of our analysis.

Correction for the impact of deforestation, forest fire, and land use change on the 11 floristic regions and areas of high botanical diversity revealed that 36% of Borneo's total surface, and 57% of its most diverse areas has already been lost. Especially the most diverse lowland forests have been severely hit by deforestation. Even more worrying is the fact that deforestation has taken its toll in IUCN recognized protected areas. Only 0.6% of Borneo's most diverse areas have an IUCN protected status, of which 33% was already deforested by the year 2000. Most dramatic is the loss of 84% of East Kalimantan's protected lowland rain forests.

To safeguard Borneo's genetic botanical diversity we urge governments and policy makers to award the remaining forested extents of areas with the highest botanical diversity with a protection status, to enforce the protection of recognized protected areas, and to conserve significant parts of each of the 11 floristic regions.

Table of Contents

CHAPTER 1
General Introduction 10

CHAPTER 2
Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images 18

CHAPTER 3
A null-model for significance testing of presence-only species distribution models 24

CHAPTER 4
Botanical richness and endemicity patterns of Borneo derived from species distribution models 38

CHAPTER 5
The floristic regions of Borneo inferred from species distribution models 56

CHAPTER 6
Borneo's remaining forests – Where to from here? 76

REFERENCES 84

Appendix Figures 94

Appendix Tables 101

Samenvatting (In Dutch) 132

Curriculum Vitae 138

Acknowledgements 140

CHAPTER 1

General Introduction

Niels Raes

Partly published as: Raes, N. and van Welzen, P.C. 2009. The demarcation and internal division of Flora Malesiana: 1857 – Present. Proceedings of the 7th International Flora Malesiana Symposium. Blumea 54(1)

BORNEO AND THE MALESIAN FLORISTIC REGION

The island of Borneo straddles the equator between latitudes 7° N and 4° S, and belongs, together with Amazonia and New Guinea, to the botanically most diverse terrestrial regions on earth (Myers et al., 2000; Barthlott et al., 2005; Kier et al., 2005). Borneo is part of the Malesian floristic region, first recognized by the Swiss botanist and explorer Heinrich Zollinger in 1857 (Zollinger, 1857; Johns, 1995). Zollinger comments on the demarcation of the 'Flora of the Dutch Indies' by F.A.W. Miquel , and argues that a floristic region should not be confused with the boundaries of a country's colonies. Based on a very limited number of distribution data and with mainly straight lines, Zollinger (1857) defined the boundaries of the Malesian floristic region (Fig 1.1; total grey area). He named his floristic region - Flora Malesiana - after the common use of the Malay language throughout the entire Archipelago. For colleagues at the time, who found the delimitation to extensive, Zollinger (1857) even recognized a 'Flora Malesiana' in a more restricted sense (Fig. 1.1; dark-grey area). He acknowledged that the western peninsular Malesia probably should be split up into three different regions: a northern-, central-, and southern region. According to Zollinger the southern region definitely belonged to the Malesian floristic region, and indeed this boundary corresponds with one of the demarcation knots of Van Steenis (1948, 1950; see below) . The reported sightings of snow-covered mountain peaks led him to conclude that the flora of New Guinea likely resembled that of a temperate mainland more than that of an island flora, hence excluded most of New Guinea from the Malesian floristic region.

Almost a century later, Van Steenis largely confirmed the Malesian floristic boundaries of Zollinger's initial delimitation, based on distribution maps of 2178 genera (van Steenis, 1948, 1950). This work was a continuation of the physiognomic map of the Dutch East Indies colonies, currently known as Indonesia, which he published in 1935 (van Steenis, 1935a, b). Van Steenis identified four contact zones and three principal 'demarcation knots' of Malesia with adjacent floral regions, viz. between the Malay Peninsula (e.g. the very south of Thailand) and Asia, between the Philippine Islands and Taiwan, the Torres Strait between New Guinea and Australia, and a less clear contact zone between the Bismarck- and Solomon islands and the Pacific islands (van Steenis, 1950). The later arbitrarily taken as eastern border because of lack of data (Fig. 1.1). The natural eastern boundary of the region lies in fact east of the Pacific Islands (van Balgooy *et al.*, 1996). It should also be noted that the demarcation knot between the Malay Peninsula and Asia is not located at the Isthmus of Kra, but through the southernmost provinces of Thailand (van Steenis, 1950). The phytogeographical status of Malesian floristic region was recently confirmed by Van Welzen *et al.* (2005), who found that 70% of 6616 sampled species was endemic to Malesia.

Wallace's Line, the Sunda Shelf, Wallacea, and the Sahul Shelf

Since the first recognition of Malesia as a floristic region, a debate is ongoing about its internal subdivision. The most famous division is in a western- and eastern sub-region, separated by Wallace's Line (Fig. 1.1; Wallace, 1860). Wallace (1860) found a distinct boundary between the Southeast Asian- and the New Guinean-Australian fauna, located east of the Philippines, between Borneo and Sulawesi and finally between Bali and Lombok. Other authors have recognized similar lines, or western and

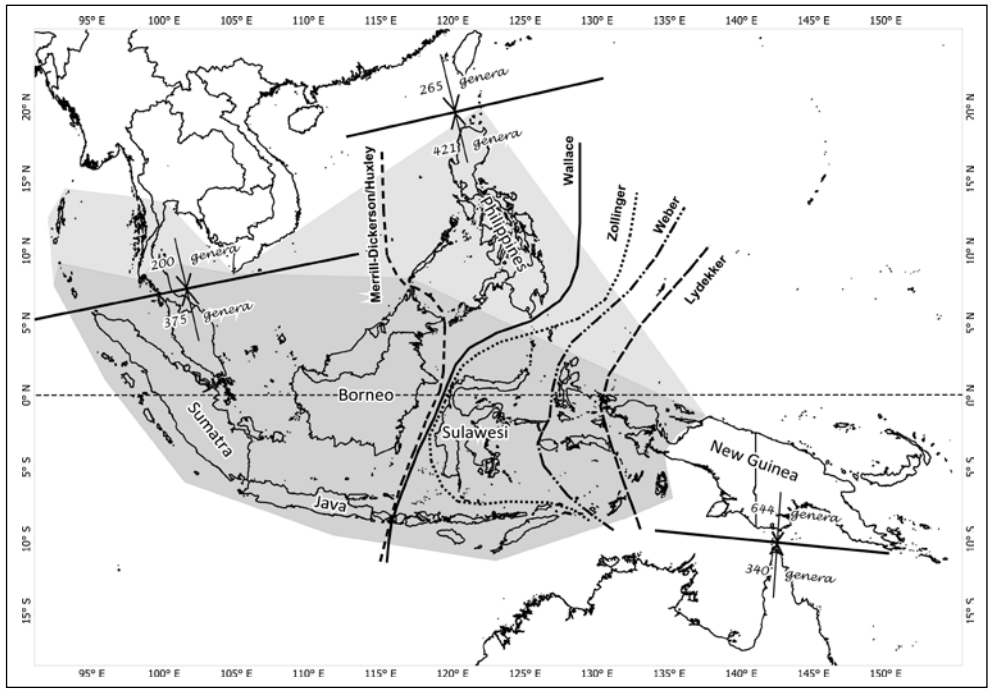


Figure 1.1. The boundaries of the Malesian floristic region defined by Zollinger (1857) in the widest sense (total grey area) and in the more restricted sense (dark-grey area); and the delimitation by Van Steenis (1948, 1950) indicated by the three demarcation knots. The numbers indicate the number of genera not crossing the knots. The different lines indicate Wallace's Line and the eastern and western variants by different authors.

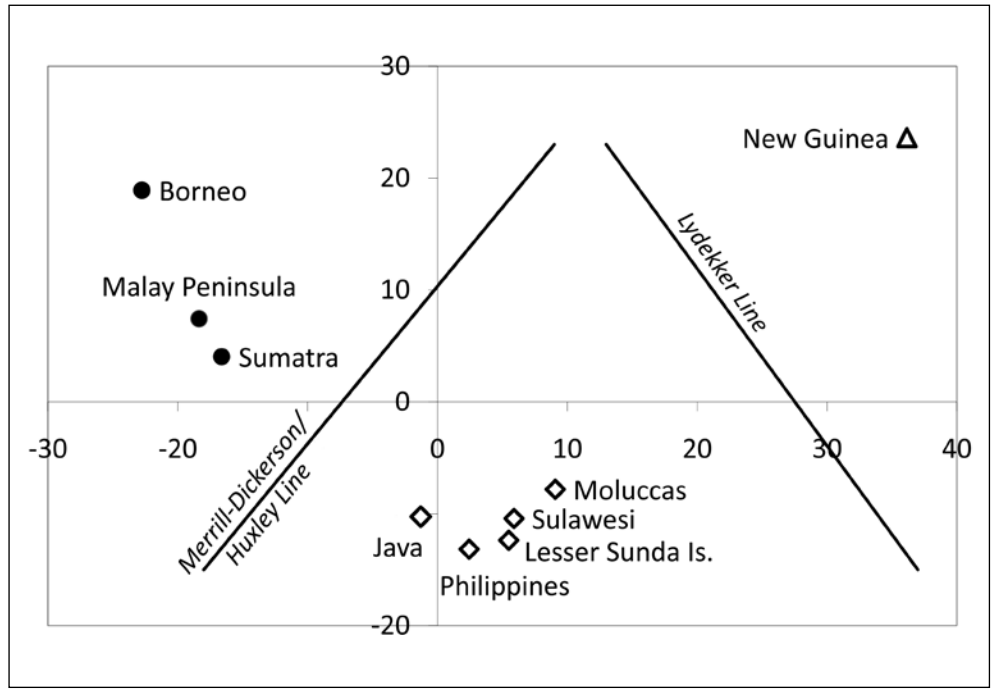


Figure 1.2. Results of the Principal Coordinate Analysis (PCO) based on presence/absence data of 7043 plant species for the nine island groups of the Malay Archipelago (from Van Welzen & Slik, 2009; Fig. 1a therein).

eastern variants of Wallace's Line (Fig. 1.1). A recent study of the evidence of the different lines based on botanical records of 6616 species showed that for all lines per side twice as many, or far more, species stop than cross the lines, and that the lines become stronger moving from west to east, meaning that less species pass a line (van Welzen *et al.*, 2005). The strong boundary of the eastern Lydekker's line indicates the very different nature of the New Guinean flora. This finding was also supported by the Principal Coordinate analysis on a slightly larger dataset containing data of 7043 species, showing the separate position of New Guinea (Fig. 1.2; van Welzen & Slik, 2009). Note that the Merrill-Dickerson/Huxley Line actually includes Java with Borneo, Sumatra and the Malay Peninsula (explained below). The floristic separation in three regions corresponds very closely with the geological history of the Malay Archipelago (Hall, 1998). The western part, west of the Merrill-Dickerson/Huxley Line in Fig. 1.1, including Borneo, Sumatra, the Malay Peninsula, and Java is also known as the Sunda Shelf. This continental shelf formed one continuous landmass during glacial maxima, when the sea levels were ~120m lower than at present, caused by an increase in land ice on the polar caps (Voris, 2000; Bird *et al.*, 2005). Under these conditions species were able to disperse to other areas on the Sunda Shelf. This has resulted in relatively high similarities in the floras of the different islands on the Sunda Shelf (Fig. 1.2; filled circles). Java is an exception, which is - contrary to the other everwet islands on the Sunda Shelf - for a large part of its surface characterized by a dry monsoon climate (van Steenis, 1979). Hence, it clusters with the islands of central Malesia, which have similar climatic conditions (Fig. 1.2; diamonds). This central Malesian region is known as Wallacea, and is located between the Merrill-Dickerson/Huxley- and Lydekker's Line, both variants of Wallace's Line. Like the continental Sunda Shelf, this central Malesian region consists of microplates which have remained submerged, only to emerge after they collided, which for Sulawesi happened only 15-10 Ma (Hall, 1997, 1998; van Welzen *et al.*, 2005). Hall (1997, 1998) provides a comprehensive overview of how the different microplates and continental platelets of Wallacea have moved, collided, emerged and submerged during the last 50 Ma. To illustrate the tectonic complexity of the Malesian region I included an image (Fig. 1.3) of one of Hall's papers (Hall, 2009). The absence of land bridges in Wallacea, disconnected the western Sunda Shelf from the eastern part of Malesia, the Sahul Shelf, also known as Papuasia (Johns, 1995). Like the Sunda Shelf, the eastern Sahul Shelf is a continental shelf which connected New Guinea to Australia during glacial maxima. The separation of the western Sunda Shelf from the eastern Sahul Shelf by Wallacea has resulted in a distinct floristic compositions on both shelves (Fig. 1.2), as is shown by Van Welzen and Slik (2009).

Botanical diversity patterns and floristic regions of Borneo

The unique status of Borneo in the Malesian region was already recognized as early as 1857 by Zollinger, who divided Malesia in five 'natural' groups, among which the 'Central-land Borneo'. He stated that the 'Central-land Borneo', in comparison to the other groups, most resembled mainland areas, and that the Malesian floristic character will likely be best expressed on Borneo (Zollinger, 1857). He also recognized that Borneo was one of the least known islands of Malesia. Since 1857 not much has changed, and Borneo with only 35 collections per 100km^2 is, after Sumatra and Sulawesi, the least collected island of Malesia (Johns, 1995). Furthermore, the Indonesian Kalimantan provinces, covering 2/3 of Borneo have the lowest collection density of the entire Malesian region with only 12 collections per 100km^2; whereas Sabah, with Mt. Kinabalu, and Sarawak together with Brunei, have 126 and 76 collections per 100km^2, respectively (Johns, 1995). This bias in collection intensities is better known as the 'Wallacean shortfall' (Whittaker *et al.*, 2005).

Nevertheless, Borneo with an estimated number of 14,423 species was found to be the most diverse island of the whole Malesian region (Roos *et al.*, 2004). A more recent analysis, based only on species treated in Flora Malesiana (Anon., 1959-2007) placed Borneo as second most diverse island, after New Guinea (van Welzen & Slik, 2009). The same study found that 37% of Borneo's vascular plant species are endemic. Furthermore, Borneo harbours four of the 'Global 200' priority ecoregions for global conservation, together covering virtually the whole island, i.e. the 'Borneo lowland and montane forests', the 'Kinabalu montane shrublands', the 'Greater Sundas Mangroves', and the 'Sundaland rivers and swamps' (Olson & Dinerstein, 2002) (See last page, Fig. a). Except for the 'Centres of Plant Diversity for Australasia' (WWF & IUCN, 1995), indicating that the centres of plant diversity on Borneo are found in smaller areas in the north, on the central mountain chain, and in the south-eastern Meratus Mountains (See last page), and a number of 'local' diversity studies (Aiba *et al.*, 2002; Potts *et al.*, 2002; Ashton, 2005; Beaman, 2005; Grytnes & Beaman, 2006) remarkable little is known about the spatial distribution of these two biodiversity components. The only study covering a larger area is the lowland Dipterocarp forests plot study of Slik *et al.* (2003).

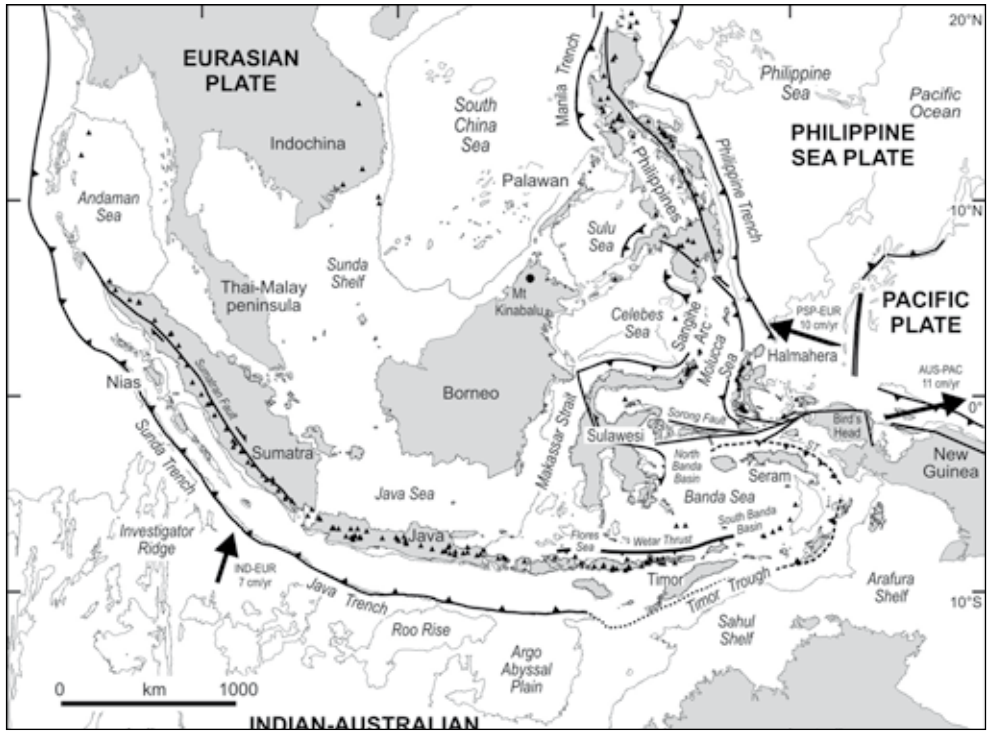


Figure 1.3. A simplified presentation of the complex present-day tectonic configuration of the Malesian region (taken from Hall, 2009; Fig. 1 therein).

The same accounts for the spatial pattern of floristic regions of Borneo. The first map delineating the different forest types (~floristic regions) of Indonesia, the former Dutch East Indies, was published in 1935 (van Steenis, 1935a). It was Van Steenis' map that served as basis for most of the following vegetation maps of Malesia (Hannibal, 1950; van Steenis, 1958b; Whitmore, 1984b; MacKinnon, 1997), ultimately resulting in the WWF 'ecoregion' map of the Indo-Pacific (Olson *et al.*, 2001; Wikramanayake *et al.*, 2002). The Bornean region of these maps is shown on the last page. Although these maps probably reflect reality to a large extent, the delineation of the floristic regions is mainly based on informal expert opinion.

Despite Borneo's exceptional botanical richness and levels of endemicity, large areas of Borneo's lowland rain forests are already deforested (Stibig *et al.*, 2007), and annual deforestation still averages 1.7% (Langner *et al.*, 2007). Even more worrying is the fact that 56% of the protected lowland forests in Kalimantan has been lost between 1985 and 2001 (Curran *et al.*, 2004). For these reasons the Sundaland hotspot, with Borneo as major component, is recognized as one of the top 5 biodiversity hotspots of the world (Myers *et al.*, 2000).

Recent digitization of the botanical collections of Borneo, housed at the National Herbarium of the Netherlands, has resulted in a database containing 166,757 records. It is this database that has provided the opportunity to quantitatively analyse the spatial patterns of botanical richness, -endemicity, and the floristic regions of Borneo, without having to rely on any informal expert opinion. Most databases containing collection records, however, suffer from a biased spatial distribution of collection records, the previously mentioned 'Wallacean Shortfall' (Whittaker *et al.*, 2005). The need to be able to predict the presence and absence of species, even for areas where no collections have been made, has resulted in a suite of species distribution modelling applications (Guisan & Zimmermann, 2000; Elith *et al.*, 2006; Peterson, 2006). Species distribution models (SDMs) predict the potential distribution of a species by describing relationships between a species' presence/absence-, or presence-only data, and a set of environmental predictors (i.e. annual precipitation, altitude, soil depth, etc.) across an area of interest, in this case Borneo. One of the remaining challenges in the field of species distribution modelling concerns the validation of SDMs developed with presence-only data, typical for herbarium collections; i.e. it is very difficult, if not impossible, to establish

the absence of a species from an area. Most measures of SDM accuracy currently applied were developed for presence/absence datasets (Fielding & Bell, 1997; McPherson *et al.*, 2004; Pearson *et al.*, 2006), and severe problems do exist when applied to presence-only data. This was also acknowledged by other authors, who placed the improvement of SDM validation high on their list of research priorities (Olden *et al.*, 2002; Guisan & Thuiller, 2005; Araújo & Guisan, 2006; Phillips *et al.*, 2006).

Borneo's botanical diversity is unique, but its threatened conservation status is of major concern. The large amount of recently digitized herbarium records, the available spatial data on global climate and soil properties, together with recent developments of species distribution modelling techniques that allow to predict the presence and absence of species even for areas that never have been sampled, make it possible to analyse the spatial patterns of botanical richness, -endemicity, and floristic regions of Borneo quantitatively at a high spatial resolution. This in turn can inform better conservation strategies for this unique natural resource.

Objectives

The objectives of this thesis are:

1. To introduce a technique known as georegistration to georeference as many collections as possible, especially for the least represented regions, to reduce the effects of collection bias to a minimum.
2. To develop a new statistical test to assess the significance of species distribution models.
3. To develop high spatial resolution botanical richness and -endemicity maps of Borneo, and to relate these patterns to environmental conditions.
4. To identify the different floristic regions of Borneo based on actual collection data, and to characterize the different regions by their environmental conditions.
5. To assess the priority regions for nature conservation on Borneo based on botanical richness, -endemicity, floristic regions and the level of deforestation.

Outline of the thesis

In Chapter 2 we introduce a technique known as georegistration. The Kalimantan provinces, in contrast to the rest of Borneo, have besides less collections also a larger proportion of the collections without coordinates required for modelling. By matching expedition maps with satellite images we attempt to georeference as many collection localities from the Kalimantan provinces as possible, thereby reducing the impact of collection bias to a minimum.
The erroneous application of the measures of model accuracy applied to presence-only species distribution models, led us to develop a new statistical significance test for this specific type of models. This method is described in chapter 3.
In chapter 4 we used all significant species distribution models to develop the botanical richness and -endemicity patterns of Borneo. The main driving factors of high levels of botanical richness and -endemicity were assessed by variance partitioning and multiple regression analyses.
In Chapter 5 we delineate the different floristic regions, based on the same significant models that were used in Chapter 4, with a hierarchical cluster analysis on the presence/absence species matrix for 8577 grid cells of Borneo. A classification and regression tree (CART) was used to characterize the ecological conditions under which each floristic region occurs.
In the final Chapter 6 we combine all results, and assess which areas of high botanical richness, -endemicity, and different floristic regions are already heavily deforested, and require most conservation efforts. Furthermore, we make suggestions for future research.

Note to the reader:
All chapters have been printed, are submitted, or are in preparation to be submitted to SCI journals. Therefore some overlap in the content of the chapters does occur.

CHAPTER 2

Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images

Niels Raes, Johan B. Mols, Luc P.M. Willemse and Erik F. Smets 2009.
Proceedings of the 7th International Flora Malesiana Symposium.
Blumea 54(1)

Abstract

For numerous scientific purposes collection records need to be georeferenced. Although the geographic coordinates of many of the collection localities are available in gazetteers, especially collections from tropical areas of the world are still not georeferenced. In an attempt to georeference these localities for Indonesian Borneo we used digitized old maps which were georegistered with SRTM digital elevation data, and Landsat 7- and JERS-1 SAR radar satellite images. This enabled us to georeference 2577 additional collections from Indonesian Borneo, belonging to 1744 taxa, which were collected at 134 previously not georeferenced localities. This applied methodology enables researchers to georeference their historical collections for biodiversity, biogeographical, and global climate change impact studies.

Keywords
georeferencing; georegistration; historical map; Landsat; JERS-1 SAR; SRTM digital elevation data

Introduction

One of the most important aspects of digitized herbarium- and natural history museum records in order to be used for i.e. biodiversity assessments, predicting the effects of -habitat loss, -potential for species' invasions, and -climate change effects (Graham *et al.*, 2004; Peterson, 2006), is that they need to be accurately georeferenced. Most collections made during the last two decades have coordinates taken with GPS equipment. The older collections, and notably those made in the 19th and early 20th century, often have only named collection localities. In order to make these older collections useful for floristic- and biogeographical research, the collection localities need to be georeferenced with the aid of a printed-, or one of the many online gazetteers (i.e. Alexandria Digital Library Gazetteer[1], La Tierra gazetteer[2], or BioGeomancer[3]). This works fine as long as the localities refer to rivers, mountains, villages etc. in western countries.

For many localities, such as small settlements, creeks, and hills in remote tropical areas, however, the coordinates have either never been assessed, or have not been made available in a gazetteer. For the purposes mentioned above, especially the collections made in remote areas can be very important, since these areas have often been visited only once by a collecting expedition. Complicating matters even further is the fact that the named localities on the labels of the collections gathered during the 19th, and early 20th century expeditions, regularly refer to vernacular names used by local guides at that time. Frequently these localities are currently known under a different name, which makes it impossible to find them in a gazetteer. Furthermore, these remote areas are likely to suffer most from the 'Wallacean Shortfall' (Whittaker *et al.*, 2005), a phenomenon that certain geographical regions are far less sampled than others, resulting in biased collection densities (Parnell *et al.*, 2003; Reddy & Davalos, 2003; Moerman & Estabrook, 2006; Hortal *et al.*, 2007). To reduce the impact of the 'Wallacean Shortfall' to a minimum, it is important to georeference as many collections as possible from these already under-collected areas. Fortunately, during the early expeditions often maps were made that indicate the collection localities and their corresponding names used at the time.

1 http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp
2 http://www.tutiempo.net/Tierra/
3 http://www.biogeomancer.org

These maps are generally stored in the very same institutions that harbour the collections.

Instead of trying to calculate the coordinates of collection localities with a ruler, based on map coordinates printed in the margins, we aimed at geographically positioning digitized expedition maps by matching them with SRTM digital elevation data and high-resolution satellite images in a geographic information system (GIS), through a process known as georegistration.

Methods

This study is part of the assessment of the botanical diversity, -endemicity, and floristic regions of Borneo derived from species distribution models (Raes & ter Steege, 2007 - Chapter 3), hence this island was used as the model. The northern and western parts of Borneo belong to the countries Malaysia and Brunei and cover 27.5 % of the total area; the remainder – the Kalimantan provinces – belong to the country of Indonesia (Fig. 2.1). Malaysia and Brunei have a long history of botanical collecting, and local biodiversity studies (Proctor *et al.*, 1983; Proctor *et al.*, 1988; Ashton & Hall, 1992; Aiba *et al.*, 2002; Potts *et al.*, 2002; Slik *et al.*, 2003; Ashton, 2005). Therefore, many collection localities of these countries have been georeferenced, and are available in a printed-, or online gazetteer. From the total of 166,757 digitized collections of Borneo present in the database of the National Herbarium of the Netherlands (NHN), 69.6 % was collected in Malaysia and Brunei. This makes it even more important to include as many georeferenced collections from the Indonesian Kalimantan provinces as possible, in order to reduce the effects of the 'Wallacean Shortfall' to a inimum. Especially for Indonesian Borneo – with its extensive network of rivers and creeks running between mountains and hills, with many small settlements along their banks – localities often have only local names which were never georeferenced. Fortunately, there exist a reasonable amount of detailed and published expedition maps from the 19th, and early 20th century (Table 2.1). These maps were used the retrieve the coordinates for as many collection localities of the Kalimantan provinces as possible.

Georegistration of digitized maps and georeferencing collection localities

The first step in the georegistration process is to digitize all available maps at high resolution (Table 2.1). Secondly, we downloaded the SRTM 90m resolution digital elevation data[1], and the 28.5m resolution Landsat 7 (circa 2000)[2] images of Borneo. The 100m resolution JERS-1 SAR[3] radar satellite images were obtained from DVD-ROM (free of charge). These data were imported in a geographical information system, Manifold GIS (Manifold Net Ltd.), and projected to a geographic projection (Kennedy, 2000).

Thirdly, the digitized maps were georegistered in Manifold GIS. Georegistration is the process of adjusting an image (the digitized maps) to the geographic location of a 'known good' reference image (the geographically projected satellite images, and SRTM digital elevation data). The georegistration process starts with the identification of one 'known good' reference feature, i.e., a major city, main river, mountain top, or an extrusion of the coast line with a (online) gazetteer. This reference point is marked on both the satellite image and the digitized map, based on the coordinates retrieved from the gazetteer. This gives an indication about the geographical position of the map, and the area it covers.

1 http://srtm.csi.cgiar.org/
2 https://zulu.ssc.nasa.gov/mrsid/
3 http://www.eorc.jaxa.jp/JERS-1/GFMP/#SEA2/

Table 2.1. List of georegistered digitized maps and their references.

Source		Reference
Geological explorations in Central Borneo		Molengraaff, 1900
	Topographic map of the north-eastern part of West Kalimantan (Map I)	
	Geological map of western Central Kalimantan and part of South Kalimantan (Map II)	
	Geological sketch-map of a part of the Kapoewas-river basin and the great lakes (Map III)	
	The Soengai Embaloeh (Map V)	
	The Soengai Mandai (Map VI)	
	The Upper Kapoewas (Map VIIa)	
	The Upper-Kapoewas, the Boengan, the Boelit and the track from the Boelit-river across the waterparting to the Mahakam-Basin in East-Borneo (Maps VIIb,c)	
	The Seberoewang and the Embahoe (Map VIIIa)	
	The Seberoewang (Map VIIIb)	
	From the Boenoet, the Sebilit and the Tebaoeng across the Madi-Plateau to the Melawi-Valley, the Lekawai and the Schwaner-Mountains (Map IXa)	
	The Boenoet (Maps IXb,c)	
	Topographical and geological sketch-map of the Samba River (Maps Xa-e)	
Comprehensive atlas of the Netherlands East Indies		van Diessen et al., 2004
	West Kalimantan pp. 350-351	
	Central and West Kalimantan pp. 352-353	
	East Kalimantan pp. 360-361	
	South and East Kalimantan pp. 362-363	
Miscellaneous		
	Banjermasing, Martapoera and part of the Lawoet areas 1845	Müller, 1857
	Kaart van de kust- en binnenlanden van Banjermasing	
	West Kalimantan	Hallier, 1895
	Sketch-map of the upper Barito (Boesang and Bakaäng) at the watersheds of the Barito-Mahakam, the Mahakam-Kapoeas and the Kapoeas-Barito.	Stolk, 1907
	Sketch-map of the Kajan, Bahau and Poedjoengan	van Walcheren, 1907
	Sketch-map of the Boeloengan and the Apo-Kajan	Nieuwenhuis, 1910
	Expeditie N.O. Borneo 1925. Reisroute v/d botanist F.H. Endert	Buys et al., 1927
	Midden-Oost-Borneo-Expeditie 1925; Endert F.H.	
	Overzicht van de tot dusverre verkregen topografische resultaten	
	Map I. Travels in the Serawai area	Winkler, 1927
	Map II. Travels in the upper Kapuas area	
	Along the Mahakam	Witkamp, 1932
	Sankoelirang	Endert, 1933
	Reede van Singkawang	Dunselman, 1939
	West Kalimantan	Dalton, 1978
	Mahakam river	
	Danau Sentarum Nature Reserve- West Kalimantan	van Balen, 1996
	Sketch-map of central East Borneo	Unknown

Next, as many reference points that were indicated on the digitized map (i.e. villages, river bends, tributaries, hill- and mountain tops), and also are recognizable on the satellite image, were marked on both the digitized map and the satellite image. Most frequently we used the Landsat 7 images,

because these have the highest resolution and the most detail. However, when a location on the digitized map was obscured by a cloud cover on the Landsat 7 satellite image, we switched to the JERS-1 SAR radar satellite image, which penetrates through the cloud cover. To identify mountain tops we used the SRTM 90m resolution digital elevation data.

Finally, the digitized map is superimposed on the satellite image based on the reference points on the satellite image, thus is georegistered (Fig. 2.1). We allowed a certain degree of transformation of the digitized maps during the georegistration process to correct for differences in map projections, i.e. the way the round earth is flattened (Kennedy, 2000), and to overcome geographical measurement errors. Remind that most of the digitized maps are originally more than a century old, and that the equipment used at the time was not as accurate as the GPS equipment used today. To georeference the remainder of the localities that were not used as reference points, we superimposed the digitized and georegistered map (set as transparent) on the satellite images. By adding the remaining localities as points on a new data layer in the GIS, we were able to retrieve their coordinates, and thereby georeferenced them. This process was repeated for all available digitized maps at the NHN-Leiden University branch (Table 2.1). The named localities with their corresponding georeferenced coordinates were exported to a spreadsheet file and merged in the Borneo gazetteer of the NHN database.

Results and discussion

In total we used 34 digitized maps. From the selection of maps shown in Figure 2.1 it is clear that they differ greatly in the area they cover, and thereby in their amount of detail. The extent to which the maps are presented as diamond shapes, instead of rectangles, indicates the accuracy of the original maps, the differences in map projections, and the degree of transformation required to match the digitized maps with the satellite images. It should be kept in mind, however, that these maps, in many cases were developed based on compass readings. Nonetheless, they were often very accurate, and allowed us to georeference many map features. It is often argued that rivers are unreliable reference points, because they change their course during time. Our georegistration experiences confirm this fact, nevertheless the ancient river bends were in many occasions clearly visible as oxbow lakes, which were regularly used as reference points in the georegistration process.
In total we georeferenced 3269 unique localities from the digitized maps listed in Table 2.1, and merged these with the Borneo gazetteer of the NHN database. These localities are represented by black and white dots in Figure 2.1. From the 50,067 (30.1%) digitized collections from Indonesian Borneo stored at the NHN, we were able to georeference 40,646 records (81%) using various sources. Of these 40,646 records, 2577 (6.34 %) were georeferenced with localities retrieved from the digitized maps. These records could be attributed to 134 unique named localities and are represented as white dots in Figure 2.1. While this is only 4.10% from the total of 3269 georeferenced unique localities, the additionally 2577 georeferenced collections represented 1744 unique taxa. Although this percentage is lower than we initially had anticipated, considering the much lower collection density of the Indonesian part of Borneo, any additionally georeferenced collections make a valuable contribution, and reduce the impact of the 'Wallacean Shortfall' to a minimum.

At the same time the additions to the Borneo gazetteer can be used by other researchers

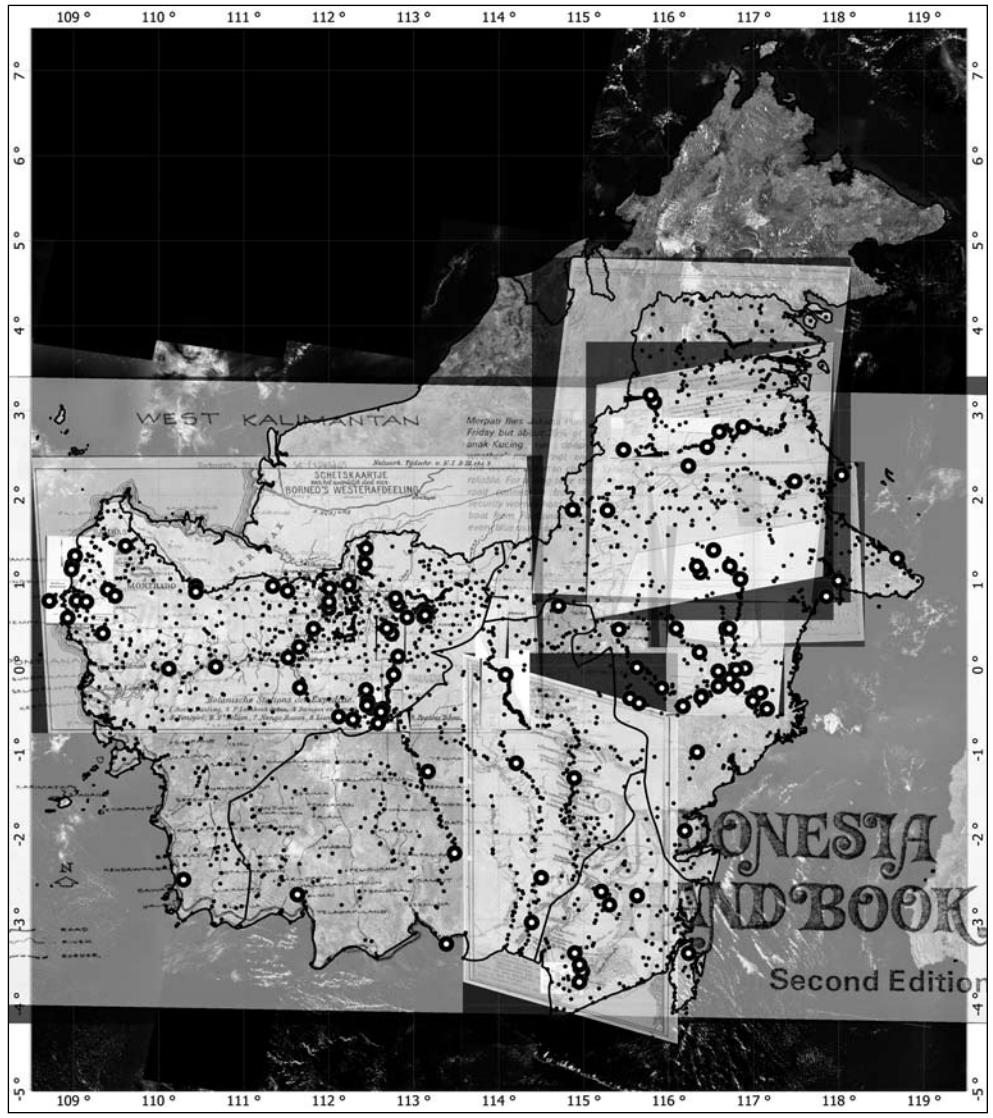


Figure 2.1. Landsat 7 image of Borneo (geographic projection) superimposed with a selection of georegistered digitized maps. Black dots indicate georeferenced localities retrieved from the maps. White dots indicate georeferenced localities where actual collections were made which otherwise could not have been georeferenced.

enabling them to georeference the records of their taxa of interest. The methodology of georegistration allows researchers to assign accurate coordinates to their specimens based on historical maps, while at the same time illustrating the importance of historical maps for current research themes.

Acknowledgements - We like to thank Ben Kieft for digitizing all the maps, Bart Meganck for usefull comments to an earlier version of the manuscript, and the staff of the National Herbarium of the Netherlands for gathering all the maps from old publications and books which were used for this paper.

CHAPTER 3

A null-model for significance testing of presence-only species distribution models

Niels Raes and Hans ter Steege

Forum
Ecography 30 (2007) 727-736

Species' distribution models (SDMs) attempt to predict the potential distribution of species by interpolating identified relationships between species' presence/absence, or presence-only data on one hand, and environmental predictors on the other hand, to a geographical area of interest. Currently, they are widely applied in biogeography, conservation biology, ecology, palaeo-ecology, invasive species studies, and wildlife management (Guisan & Zimmermann, 2000; Araújo & Pearson, 2005; Thuiller *et al.*, 2005; Araújo & Guisan, 2006; Guisan *et al.*, 2006; Peterson, 2006). More recently, vast numbers of herbarium and natural history museum collections have become available (Graham *et al.*, 2004) and techniques to apply this special type of presence-only data have been developed (Hirzel *et al.*, 2002; Anderson *et al.*, 2003; Elith *et al.*, 2006; Pearce & Boyce, 2006; Phillips *et al.*, 2006). Despite the widespread use of SDMs, several high-priority research interests remain to be investigated (Guisan & Thuiller, 2005; Araújo & Guisan, 2006). One of these is the improvement of SDM validation, or the quantification of a model's predictive performance (Araújo & Guisan, 2006).

The fact that the standard validation procedures for an SDM are not sufficient to assess the applicability of an SDM in a predictive context, was first shown by Olden *et al.* (2002). They showed that after SDM validation it is critical to assess whether the SDM prediction differs from what would be expected on the basis of chance alone. SDMs producing random predictions are neither helpful nor useful (Olden *et al.*, 2002). Thus, in this paper we introduce a null-model methodology that allows testing whether SDMs developed with presence-only data differ significantly from what would be expected by chance. We also demonstrate that it is critical and possible to correct for collector-bias in specimen data in this test.

SDM validation and measures of accuracy

Validation of SDMs can be carried out with several different measures of model accuracy. The most widely applied measures of model accuracy include sensitivity, specificity, Cohen's kappa, and the area under the curve (AUC) of the receiver operating characteristic (ROC) plot (Fielding & Bell, 1997; Manel *et al.*, 2001; McPherson *et al.*, 2004). Most measures of SDM accuracy, including the four mentioned above, are directly or indirectly derived from a confusion matrix (see Fielding and Bell 1997). Sensitivity quantifies the proportion of observed presences correctly predicted as presence, the true positive fraction. Specificity quantifies the true negative fraction. Cohen's kappa quantifies overall agreement between predictions and observations, corrected for agreement expected to occur by chance.

These three measures of accuracy require that probabilities of occurrence obtained with SDMs are transformed into discrete presences or absences, for which purpose a threshold of 0.5 is commonly used (McPherson *et al.*, 2004; Liu *et al.*, 2005; Jiménez-Valverde & Lobo, 2007).

The AUC value of the ROC plot is a method that does not require discrete presence/absence predictions, and is therefore a measure of accuracy that is threshold independent (Pearce & Ferrier, 2000; McPherson *et al.*, 2004).

The ROC plot is obtained by plotting sensitivity as a function of the falsely-predicted positive fraction, or commission error (1-specificity), for all possible thresholds of a probabilistic prediction of occurrence. The resulting area under the ROC curve provides a single measure of overall model accuracy, which is independent of a particular threshold. AUC values range from 0 to 1, with a value of 0.5 indicating model accuracy not better than random, and a value of 1.0 indicating perfect

model fit (Fielding & Bell, 1997). An AUC value can be interpreted as indicating the probability that, when a presence site (site where a species is recorded as present) and an absence site (site where a species is recorded as absent) are drawn at random from the population, the presence site has a higher predicted value than the absence site (Elith *et al.*, 2006; Phillips *et al.*, 2006).

All four measures of model accuracy were tested extensively for statistical artefacts, and the AUC value was the only measure of SDM accuracy that was invariable to the proportion of the data representing species' presence, known as prevalence (Pearce & Ferrier, 2000; Manel *et al.*, 2001; McPherson *et al.*, 2004). Insensitivity to prevalence is of special relevance when the AUC values are used to assess model accuracy for SDMs that have been developed with presence-only data. When the required absences are lacking, they are replaced by pseudo-absences. Pseudo-absences are sites, randomly selected across the geographical area of interest, at localities where no species presence was recorded and for which species occurrence is set as absent (Ferrier *et al.*, 2002; Anderson *et al.*, 2003; Elith *et al.*, 2006; Phillips *et al.*, 2006). A sufficiently large sample of pseudo-absences is needed to provide a reasonable representation of the environmental variation exhibited by the geographical area of interest, typically 1,000-10,000 points (Stockwell & Peters, 1999; Ferrier *et al.*, 2002; Phillips *et al.*, 2006). These large numbers of pseudo-absences automatically result in low prevalence values. The number of records by which a species is represented in herbaria and natural history museums range from one to 150-200 records (Stockwell & Peterson, 2002). Even when a species is represented by 200 unique presence-only records and 1,000 pseudo-absences are used, prevalence is only 16.7% (200/1200).

A major drawback of using pseudo-absences, however, is that the maximum achievable AUC value indicating perfect model fit, is no longer 1, but 1-*a*/2 (where *a* is the fraction of the geographical area of interest covered by a species' true distribution, which typically is not known (Phillips *et al.*, 2004; Phillips *et al.*, 2006). Nevertheless, random prediction still corresponds to an AUC value of 0.5. Therefore, standard thresholds of AUC values indicating SDM accuracy (e.g., the threshold of AUC>0.7 that is often used; Pearce and Ferrier 2000, Swets *et al.,* 2000, Manel et al. 2001), do not apply.

A null-model approach for significance testing of presence-only SDMs

To test the significance of an SDM we propose to test the AUC value (of the SDM) against a null distribution of expected AUC values based on random collection data (sensu Olden *et al.* 2002). A null-distribution, or null-model, is a model that is based on randomizations of ecological data or random sampling from a known or imagined distribution (Swets *et al.*, 2000; Jetz *et al.*, 2004; Gotelli & McGill, 2006). A null-model is straightforward in theory and closely resembles hypothesis testing in conventional statistical analysis. To build a null-model, first the AUC value of the real SDM is determined. Next, a null-model is generated by randomly drawing collection localities without replacement, from the geographical area for which the species distribution is modelled. The number of randomly drawn collection localities is equal to the actual number of collections for that species. This is repeated 999 times to generate a frequency histogram of AUC values, expected if the null hypothesis is true. The position of the observed AUC value in the null distribution of the 'randomly' generated AUC values is then used to assign a probability value, just as in a conventional statistical analysis (Dolédec *et al.*, 2000; Olden *et al.*, 2002; Gotelli & McGill, 2006). We use a one-sided 95% confidence interval (C.I.) since we are only interested in whether an SDM performs significantly better than expected by chance, rather than assessing whether it performs significantly worse. We interpret a significant model to indicate that the relations between species' presence localities and the predictor variable values at those locations are stronger than can be expected by chance.

An additional advantage of significance testing of an SDM with a null-model is that we can use all presence records to develop and test the SDM. Common practice in measuring an SDM's accuracy is the split-sample approach. This approach splits the available species records into a training and test sample (Fielding & Bell, 1997). It is assumed that a randomly selected test sample from original data constitutes independent observations, which can be used for statistical testing (Araújo *et al.*, 2005). However, such a test sample is not fully independent due to spatial autocorrelation (Araújo *et al.*, 2005; McPherson & Jetz, 2007). Moreover, dependent on the random split, different values of SDM accuracy may be obtained (Phillips *et al.*, 2006). Phillips *et al.* (2006) showed that SDMs for a species represented by 128 records and 10 different random splits, yielded AUC values ranging from 0.819 until 0.903. More extremely, our unpublished results yielded AUC values for a species represented by 8 records ranging between 0.079 and 0.912 based on 100 random splits.

Testing an SDM against a null-model, however, could suffer from one more problem. When drawing random points from a geographical area one assumes that collectors visited all localities equally well. If this condition is not met, which is likely to be the case (Reddy & Davalos, 2003; Romo *et al.*, 2006; Hortal *et al.*, 2007), the randomly drawn points, that are used to develop the null-model, might include ecological conditions that are not represented by the localities from where actual collections were gathered. This bias could results in a significant deviation from the null-model for species that are randomly distributed over the actual collection localities.

The impact of collection bias on significance testing

SDMs predict the presence and absence of a species for a given geographical area, based on the localities where the records were collected and the values of environmental predictors at those sites. SDMs are especially useful when only part of the entire geographical area has been sampled, as is generally the case. This works fine as long as the collection localities are randomly spread over the complete geographical area. Unfortunately, collectors tend to visit areas which are easily accessible, such as areas close to cities, roads, rivers, and nature reserves resulting in serious collection biases (Parnell *et al.*, 2003; Reddy & Davalos, 2003; Kadmon *et al.*, 2004; Hortal *et al.*, 2007). The influence of collection biases on the accuracy of SDMs largely depends on the range of values of each of the environmental variables covered by the collection localities, known as climatic, or environmental bias (Kadmon *et al.*, 2003, 2004). Kadmon *et al.* (2003) showed that environmental biases, expressed as the degree of sampling bias

with respect to the environmental conditions under which a species is known to occur, had a significant negative effect on the predictive accuracy of the SDM. Although this is a serious issue of concern (Araújo & Guisan, 2006), it is not specific to any methodology used to develop SDMs. However, it is relevant when the accuracy of an SDM is tested against a null-model.

When collecting is environmentally biased, an SDM is more likely to deviate significantly from a random null-model that does not include such bias. When, for example, collection localities are biased for mean annual temperature, a significant part of the species' actual temperature range could remain unsampled. When these data are used in an SDM that is tested against a null-model, based on records that were randomly drawn from the entire study area, this species will possibly show a preferred mean annual temperature range compared to the randomly drawn points. It will accordingly more likely deviate significantly from the null-model than its actual range would justify. Such collection bias might thus result in certain areas being systematically under predicted by the SDM. It should be noted, however, that this is true for all distribution modelling methods and can only be solved by additional data collection.

Fortunately, the problem of having a higher chance of significantly deviating from a randomly drawn null-model if collections are biased, can be solved by restricting the randomly drawn points to all known collection localities. Thus, drawing the null-model from a biased distribution. To test for environmental bias in known collection localities a distribution model using all known collection localities is tested against a null-model developed by 100 -1000 times drawing an equal number of random points from the entire study area. If the distribution model's accuracy of known collection localities deviates significantly from this 'second' null-model, then we conclude that the collection localities are environmentally biased. If this is the case then the SDMs have to be tested against a null-model that is based on actual collection localities.

A case study based on Bornean plant collections

To illustrate the applicability of a null-model approach to select SDMs that deviate significantly from random expectation, we selected all occurrences of the genus *Shorea* (Dipterocarpaceae) on the Malesian island Borneo (approx. 8°N - 5°S, 108° - 120°E; Fig. 3.3) from the BRAHMS database of plant collections present at the National Herbarium of the Netherlands, Leiden University, the Netherlands. *Shorea* was selected because this genus has been thoroughly taxonomically revised and species identifications are reliable (Ashton, 1983). The database contained 4466 records of 147 Shorea species for Borneo. Out of these 147 species, 116 were represented by 5, or more, unique collection localities. For those species, we developed SDMs.

To model the species distributions we used environmental predictor variables with a 5 arc-minutes resolution (~10km at the equator). We selected the digital elevation model (DEM) and the 19 bioclimatic variables of the current conditions (~1950-2000) from the WORLDCLIM dataset (hhtp://www.worldclim.org) for Borneo (Hijmans *et al.*, 2005). Additionally, we selected 15 FAO soil variables (FAO, 2002). We also included a measure of the effect of the El Niño Southern Oscillation Event (ENSO). This variable was expressed as the relative average annual difference in Normalized Difference Vegetation Index (NDVI) between the months of an ENSO, and a non-ENSO year. To this dataset

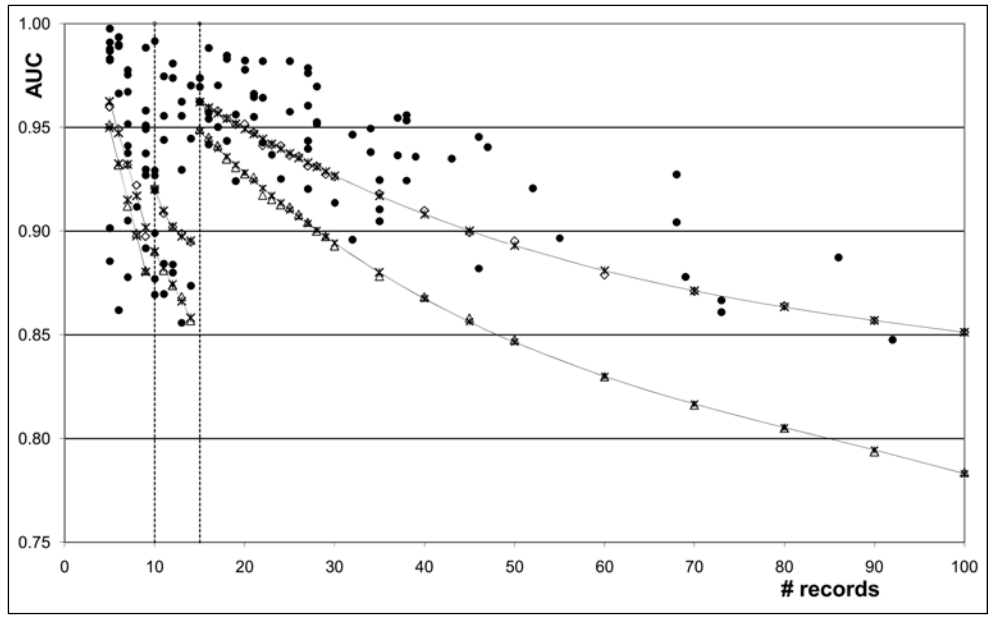


Figure 3.1. Species' distribution model (SDM) AUC values (•), the 95% confidence interval (C.I.) AUC values of the randomly drawn null-models (Δ), and the 95% C.I. AUC values of the environmentally bias corrected null-models (◊). Asterisks give the fitted 95% C.I. AUC values for both series of null-models connected by a line. Vertical dotted lines indicated the consecutive addition to the initial linear modelling features, of quadratic, and hinge features by Maxent. SDM AUC values that are higher than their corresponding 95% C.I. AUC value of the fitted null-model, significantly deviate from what would be expected by random chance ($p<0.05$).

we added the Walsh's index (Walsh, 1996; Leigh Jr., 2004). This index integrates the effects of annual rainfall and its seasonality. Finally, the elevation range derived from the SRTM 90m Digital Elevation Data (http://srtm.csi.cgiar.org/) was added. All data layers were scaled to 5 arc-minute resolution, and resampled to the geographical extent of the most restricted FAO soil variable data layers. This resulted in 8577 data cells for Borneo. All data layer manipulations were performed with Manifold GIS (Manifold Net Ltd).

To model *Shorea* species distributions of Borneo we used Maxent (version 2.3.0; http://www.cs.princeton.edu/~shapire/maxent/) (Phillips *et al.*, 2006). Maxent, or the maximum entropy method for species' distribution modelling, estimates the most uniform distribution ("maximum entropy") across the study area, given the constraint that the expected value of each environmental predictor variable under this estimated distribution matches its empirical average (average values for the set of species' presence records) (Hernandez *et al.*, 2006; Phillips *et al.*, 2006). Maxent was specifically developed to model species distributions with presence-only data and has outperformed most other modelling applications (Elith *et al.*, 2006; Hernandez *et al.*, 2006; Pearson *et al.*, 2007). An added advantage of Maxent is that it also performs the ROC statistical analysis. Since we tested whether an SDM's AUC value deviates significantly from a null-model, the 'random test percentage' was set to zero resulting in training data only. To avoid the inclusion of multiple presence records in one grid cell per species we set Maxent to 'remove duplicate presence records'. This reduced the total available presence records for the 116 *Shorea* species represented

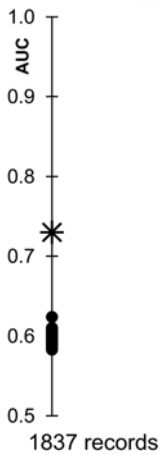


Figure 3.2. The AUC value of the model based on the 1837 collection cells (∗) and the 100 AUC values (•) of models based on 1837 randomly drawn cells from the total 8577 cells of Borneo, indication the 1837 collection cells are significantly environmentally biased (p<0.01).

by at least five records to 2552. The modelling rules were set to 'Auto Features' using only linear features when less than 10 records were available, adding quadratic features for SDMs developed with 10 or more and less than 15 records, and including hinge features for species with 15 or more records. Maxent adds product and threshold features for those species represented by 80, or more, records. However, we set Maxent to use linear, quadratic and hinge features for all species represented by at least 15 records, due to odd behaviour of Maxent when product and threshold features were added (explained in the discussion). For each of the 116 *Shorea* species we developed an SDM with Maxent using all presence records under the modelling rules as described above. The number of unique records per species ranged from 5 until 92 (Table S3.1, '# records'). The AUC values of all *Shorea* SDMs are presented as dots in Figure 3.1, and under 'AUC' in Table S3.1.

Testing SDMs against a null-model

To test whether *Shorea* SDMs significantly differed from what would be expected by chance, we calculated the 95% C.I. AUC value for each number of records by which the *Shorea* species were represented. We developed frequency histograms of expected AUC values by randomly drawing points without replacement from all 8577 available cells of Borneo (999 times), and model these with Maxent under the same conditions as the *Shorea* species. We developed frequency histograms of expected AUC values for 5 – 30 records (26 distributions), for 35 – 50 records with intervals of 5 records (4 distributions), and for 60 – 100 records with intervals of 10 records (5 distributions). For each frequency histogram, we assessed the 95% C.I. upper limit AUC value, by ranking the 999 AUC values and selecting the 949th value (0.95 x 999 = 949; Fig. 3.1, triangles). For each of the three resulting sets of 95% C.I. AUC values we applied a curve-fit (Fig. 3.1, asterisks). The fitted 95% C.I. AUC values of the null-models for the number of records by which each *Shorea* species is represented, are given in Table S3.1, '95% C.I. All'.

With the fitted 95% C.I. AUC values, it is now easy to assess which of the *Shorea* species has an accuracy of its SDM that is significantly higher than expected by chance alone ($p<0.05$). This was the case for 105 of the 116 *Shorea* species (91%) which were modelled (Table S3.1, '95% C.I. All').

Testing SDMs against a bias corrected null-model

In order to assess whether the known collection localities are environmentally biased, we selected all databased and georeferenced plant specimen records from Borneo that were present in the BRAHMS database of the National Herbarium of the Netherlands. In total the database contained 142,097 properly georeferenced records. These records could be assigned to 1837 of the total of 8577 grid cells of Borneo This means that only 21.4% of the grid cells of Borneo have been

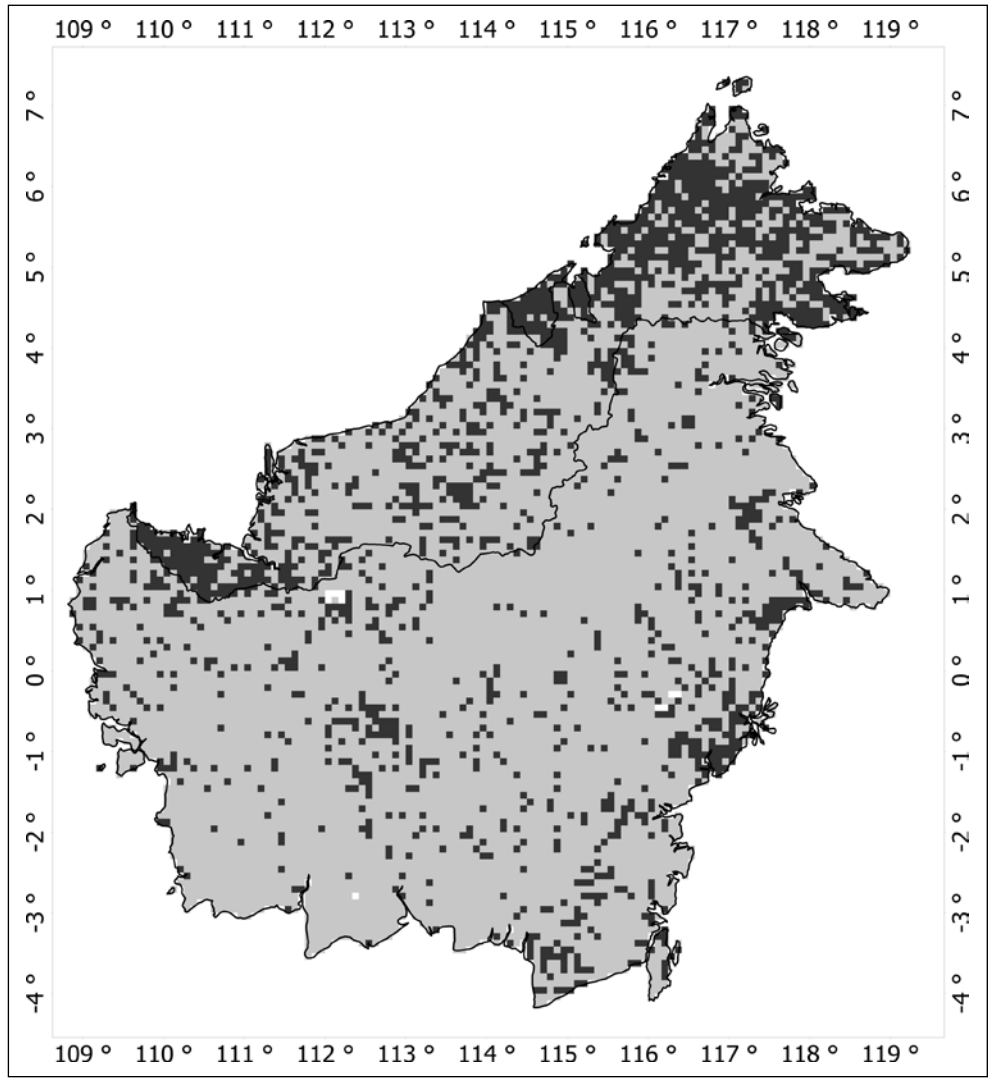


Figure 3.3. Spatial distribution of the 1837 cells, from the 8577 cells for Borneo, where at least one of the 142 097 collections was made (indicated by dark grey squares). Light grey squares indicate the remaining 6740 unsampled cells. White cells indicate large lake areas for which no environmental data were available.

visited by collectors who actually made any collections (Fig. 3.3). The collections are clearly geographically biased, as evident from the geographical distribution of the dark grey squares in Figure 3.3. However, predicting species presences or absences in non-visited areas is one of the major applications of the use of SDMs, so this should not be a major problem. More importantly, it is to assess whether these localities are environmentally biased, or whether certain conditions are over- or under-represented with respect to the environmental conditions for the entire geographical area of Borneo. For this purpose,

we first developed a distribution model of the 1837 collection localities and assessed the model's AUC value. Then, we developed a frequency histogram of expected AUC values on basis of 1837 randomly drawn localities from the 8577 cells of Borneo (100 reps). Unfortunately the AUC value of the distribution model based on the collections localities, is significantly different from random expectation (p<0.01; Fig. 3.2), hence, the collection localities are also environmentally biased.
The implication that collecting effort is environmentally biased for Borneo is that SDMs cannot be tested with null-models drawn randomly from all 8577 grid cells of Borneo.
To overcome this problem we developed a second series of null-models, in the same way as described above, but now randomly drawing from the 1837 known collection locality cells. The resulting 95% C.I. AUC values of these null-models are presented as diamonds in Figure 3.1. Again, we applied a fit through these values to establish the 95% C.I. AUC values against which the SDM AUC values were tested. These values are given in Table S3.1 under '95% C.I. Bias'. Now only 80 of the 116 *Shorea* species (69%) have a SDM AUC value significantly different from a (bias corrected) null-model (Table S3.1, '95% C.I. Bias'; Fig. 3.4a,c). This means that an additional 25 SDMs were rejected, compared to testing against environmentally unbiased null-models.

Discussion

By proposing the use of null-models in the field of presence-only species' distribution modelling, we introduce a novel methodology that allows for significance testing of SDMs. The new methodology makes use of all presence records to develop an SDM and to test its accuracy with the AUC procedure, a threshold- and prevalence-independent single measure of SDM accuracy. A significant SDM indicates that correlations between species' presence localities and the environmental predictor variables, as identified and interpolated by Maxent, deviate from random chance.
Secondly, we show the importance of correcting for environmental biases in data collection. Null models which incorporate the environmental bias within the collection data reject a significant fraction of SDMs which are significant based upon a randomly drawn null-model. If the collection localities are environmentally biased and a species is found throughout the subset of values represented by the collection localities, this species is likely to differ significantly from a null-model which is drawn from the total range of values. This results in an SDM that is an underestimation of the true geographical range of the species. This, because under these conditions the full range of values under which the species truly occurs is not incorporated in the SDM.
Although we introduce a null-model approach to the field of presence-only species' distribution modelling, the use of null-models for significance testing was successfully applied by Olden *et al.* (2002) for presence-absence SDM testing, and by Dolédec *et al.* (2000) in the field of community analysis. Our methodology differs from Olden *et al.* (2002) in that we adapted the null-model approach to make use of presence-only data, and test an SDM accuracy with the threshold- and prevalence independent AUC procedure (Swets, 1988; Manel *et al.*, 2001; McPherson *et al.*, 2004; Guisan *et al.*, 2006). This is important as in our case study the number of species presence records ranged from 5 to 92. Combined with 1,000 pseudo-absences this resulted in prevalence values as low as 0.5 to 8.4%.
We interpret that species, for which the SDM AUC value significantly deviates from a null-

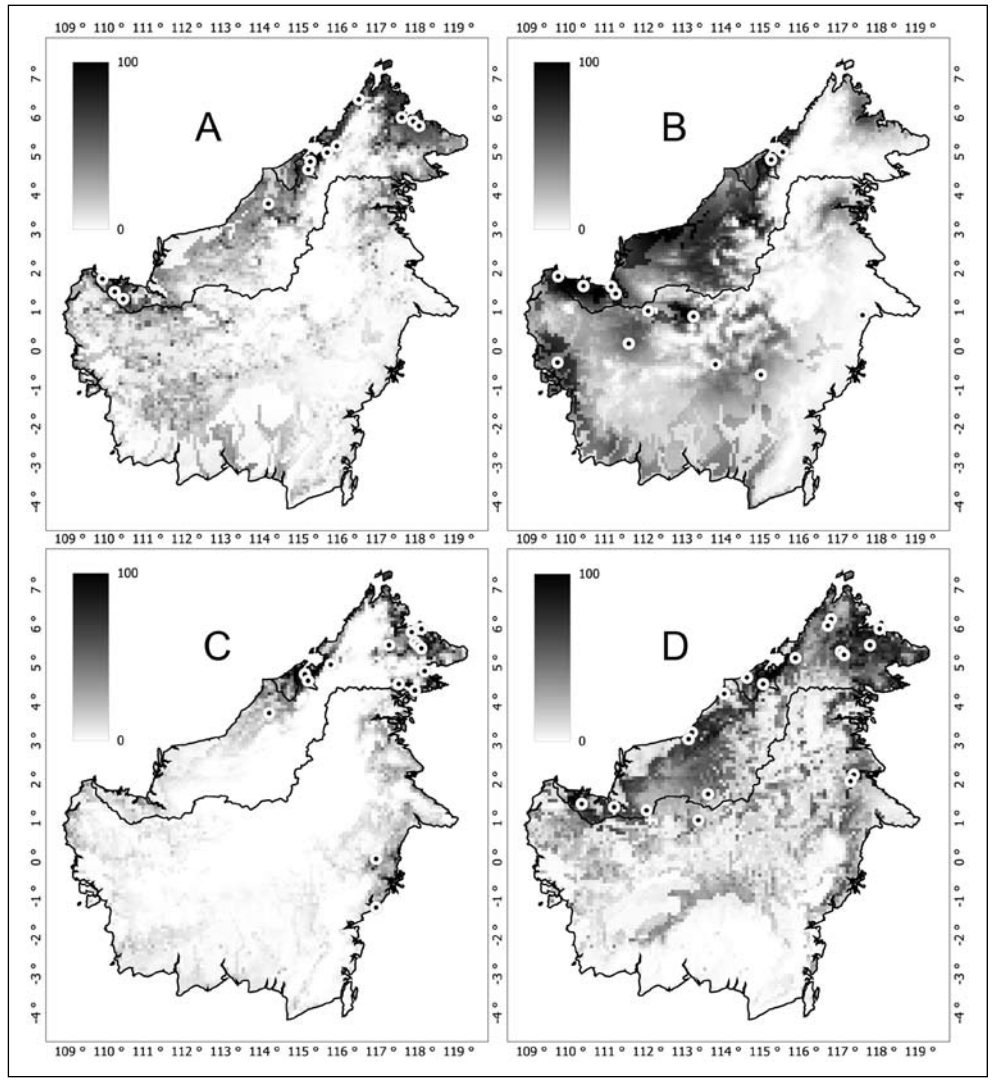


Figure 3.4a-d. Maxent predictions for two significant SDMs (A, C), and two non-significant SDMs (B, D). Collection localities are indicated by dots. A) *Shorea isoptera P.S. Ashton*, (Appendix, Table S3.1, #45), B) *S. platycarpa Heim* (Appendix, Table S3.1, # 49), C) *S. confusa P.S. Ashton* (Appendix, Table S3.1, #57), and D) *S. macroptera Dyer* (Appendix, Table S3.1, #66).

model, have specific niche requirements that were met at the localities where they were collected. This agrees with the reasoning of Dolédec *et al.* (2000). They analysed community data with a new multivariate method they called OMI (for Outlying Mean Index), to measure the distance between mean habitat conditions used by a species, and the mean habitat conditions of the sampling area (Dolédec *et al.*, 2000). The OMI value (analogous to the SDM AUC value) of a species is tested against the null-distribution of '1000 random permutation values obtained

under the null hypothesis that the species is indifferent to its environment'. For species that significantly deviated from this, 'theoretical ubiquitous species that tolerates the most general habitat conditions', it was concluded that the observed species position in habitat differed significantly from what would be expected by chance. This OMI-methodology was later implemented in a species distribution modelling technique called Ecological-Niche Factor Analysis (ENFA) (Hirzel *et al.*, 2002), but testing against a null-distribution was never formalized.

The first to notice that accuracy assessment of presence-only SDMs alone was not sufficient, and SDMs should be tested against a random null hypothesis were Anderson *et al.* (2002). They used the split sample approach, dividing the available presence records of a species in a 75% training and 25% test dataset. After SDM development using the training data, they tested whether test points fell into areas predicted presence more often than expected at random, given the overall proportion of pixels predicted presence vs. predicted absence for that species (Anderson *et al.*, 2002). The latest advances in this methodology were recently made, by introducing a jack-knife (or 'leave-one-out') procedure for SDM accuracy assessment and a combined *p*-value significance test for significance testing of the presence-only SDMs (for details see Pearson *et al.* 2007). However, this methodology does not take into account possible environmental bias in collection localities. If the full niche of a species is not represented by the collection localities, the species' predicted distribution will be smaller than its true distribution. Modelling applications, such as Maxent, are very well capable of predicting the species' distribution based on the available presence records without model under-fitting. A smaller predicted species' distribution automatically results in higher chance of significantly deviating from the random null hypothesis, the same way as in our case study more species significantly deviate from a randomly drawn null-model than from a null-model that is corrected for environmental bias (Fig. 3.1; Table S3.1). Additionally, the jack-knife validation approach may lead to overoptimistic estimates of the predictive power with larger sample sizes (Pearson *et al.*, 2007).

Our results showed the importance of correcting for environmental bias in known collection localities when null-models are used for significance testing of presence-only SDMs. However, at the same time this requirement hampers the general applicability of the methodology. In our case study, we could make use of the full herbarium record database of the National Herbarium of the Netherlands, containing 142,097 georeferenced plant specimen records found in 1837 of the 8577 grid cells of Borneo. We recognize that this amount of data will not always be available. However, since the majority of collections has been made in close proximity to roads, rivers, cities, and nature reserves (Reddy & Davalos, 2003; Kadmon *et al.*, 2004; Hortal *et al.*, 2007), an alternative could be to use a distance buffered road-river map, including cities and nature reserves, to select the grid cells and test these cells for environmental biases. If these cells are environmentally biased, the SDMs can then be tested against a null-model drawn from this pool of cells. However, this approach is less accurate and requires further testing.

Our results showed that for low prevalence values very high AUC values can be expected from randomly drawn points (Fig. 3.1; Table S3.1). Olden *et al.* (2002) too reported such high accuracy values for low (and high) prevalence. The 95% C.I. AUC value of the bias corrected null-model for 15 records (prevalence = 1.48%) was as high as 0.9622 (Fig. 3.1; Table S3.1).

Nevertheless, 80 of the 116 *Shorea* species (69%) had an SDM AUC value higher than the 95% C.I. AUC value of the bias corrected null-model. Dolédec *et al.* (2000) reported that, for their application of a null-model for two case studies, 59% and 85% of their species respectively, had significant results. Pearson ***et al.*** (2006) report values from 62-100% depending on the modelling application and thresholds that were used. Our testing against a randomly drawn null model resulted in a comparably high percentage (91%) of significant SDM AUC values ($p<0.05$). All these results are higher than the 50% reported by Olden *et al.* (2002).

Both the AUC values of the two null-models, and the SDMs, show a decreasing trend with increasing number of records (Fig. 3.1; Table S3.1). This is most likely the result of applying ROC plots to SDMs, developed with presence-only data, reducing the maximum AUC value dependent on the species' true distribution (Phillips *et al.*, 2006). Assuming that the predicted species' distributions are a good proxy for the species' true distributions, we assessed the area for which species were predicted to be present by converting the continuous probabilistic Maxent predictions of occurrence to discrete presence-absence values. We used the maximized sensitivity-specificity sum threshold for this purpose (Liu *et al.*, 2005). Regressing significant SDMs AUC values against the area for which they were predicted to be present (Table S3.1; 'Area (in %)') revealed a significant negative linear correlation (AUC=0.9913-0.0029*Area; $p<0.001$; $R^2 = 0.576$). We consider this as a strong indication that it is not the accuracy of the models that is reduced but merely that the maximum achievable AUC value is reduced due to an increased true distribution of the species. We therefore do not support the statement that the predictive accuracy of the model decreases when the extent of a species distribution increases, as suggested by Hernandez *et al.* (2006). When an increased predicted distribution and related lower SDM AUC value is caused by a broad niche amplitude, however, as is the case for habitat generalists, an SDM accuracy is more likely not to deviate from a null-model and the SDM can therefore not be used. This is possibly the case for the SDMs presented in Figure 3.4b,d (Table S3.1; #49 and #66).

A consequence of implementing the proposed use of null-models for SDM evaluation, is that SDM accuracy is tested with the same data used to develop models, i.e., a form of model verification (Araújo & Guisan, 2006). A problem with this approach is that SDMs may over-fit the calibration, or training data (Araújo *et al.*, 2005). Over-fitting, however, is not considered a problem if the goal is to describe a pattern and simultaneously reduce false negatives: i.e., true observations that are not predicted by the model (Araújo & Guisan, 2006). An advantage is that all observations are used to develop the SDMs, making optimal use of all available information. If the modelled species' distributions are intended to be used for conservation planning, verification is an approved method to test whether an SDM performs as intended. However, if the models are used to predict range shifts under different climate change scenario's, or to assess the possible invasiveness of a species, an SDM's ability to correctly predict independent test data is preferred (Araújo & Guisan, 2006). It should be kept in mind, however, that SDMs, as they are applied in this study, predict the potential distribution of a species and do not take into account competition, and historical or present geographical barriers (Soberón & Peterson, 2005; Peterson, 2006). Most studies addressing these issues use data partitioning methods to allocate records to training and test datasets. The most familiar technique is one-time data-splitting (Araújo *et al.*, 2005). Our unpublished

results indicated, however, that dependent on the spatial distribution and the random split of the records, SDM accuracies could be very different.

An advantage of Maxent is its ability to counteract the tendency of SDMs to over-fit when few presence records are available, due to its regularization procedure (Hernandez *et al.*, 2006; Phillips *et al.*, 2006). Therefore, we used the standard settings of Maxent. However, the null-models developed for 80, 90, and 100 records developed with the modelling rules set to 'auto features' and the regularization multiplier set to 1, resulted in increasing 95% C.I. AUC values indicating over-fitting of the models (data not shown). For this reason, we set the modelling rules to use linear, quadratic and hinge functions to develop the null-distributions for those numbers of records and the SDMs developed with more than 79 records.

We are aware that spatial autocorrelation in the distribution of the species records and environmental variables may also influence SDM accuracy. Our intention was not to investigate the influence of spatial autocorrelation on SDM accuracy, however, but to provide a methodology for significance testing of presence-only SDMs. Simultaneously we showed that the evaluation of presence-only SDM quality based on subjective ROC plot thresholds (e.g. AUC≥0.7 = useful), cannot be applied. With this contribution, we hope to provide SDM users with a valuable tool to identify those species that can be accurately modelled, while providing an additional reason for being cautious about interpretations of SDMs that are not tested for significance.

Acknowledgements – We would like to thank Steven Phillips for adapting the Maxent application several times to allow us to perform the ROC analyses of the null-models. We also thank Pieter Baas, Chuck Cannon, Peter Hovenkamp, Marco Roos, Ferry Slik and four anonymous referees, all of whose comments allowed us to make useful improvements to the manuscript.

CHAPTER 4

Botanical richness and endemicity patterns of Borneo derived from species distribution models

Niels Raes, Marco C. Roos, J.W. Ferry Slik, E. Emiel van Loon, and Hans ter Steege

Ecography (accepted)

Abstract

This study provides a Borneo-wide, quantitative assessment of botanical richness and endemicity at a high spatial resolution, and based on actual collection data. To overcome the bias in collection effort, and to be able to predict the presence and absence of species, even for areas where no collections have been made, we constructed species distribution models (SDMs) for all species taxonomically revised in Flora Malesiana. Species richness and endemicity maps were based on 1439 significant SDMs. Mapping of the residuals of the richness-endemicity relationship identified areas with higher levels of endemicity than can be expected on the basis of species richness, the endemicity hotspots. We were able to identify one previously unknown region of high diversity, the high mountain peaks of East Kalimantan; and two additional endemicity hotspots, the Müller Mountains and the Sangkulirang peninsula. The areas of high diversity and endemicity were characterized by a relatively small range in annual temperature, but with seasonality in temperatures within that range. Furthermore, these areas were least affected by El Niño Southern Oscillation drought events. The endemicity hotspots were found in areas, which were ecologically distinct in altitude, edaphic conditions, annual precipitation, or a combination of these factors. These results can be used to guide conservation efforts of the highly threatened forests of Borneo.

Introduction

Borneo, the third largest island of the world, is the botanically most diverse part of the Sundaland hotspot, one of the world's 25 biodiversity hotspots (Myers *et al.* 2000). Southeast Asia as a whole faces an estimated loss of three quarters of its original forest area by 2100, and up to 42% of its biodiversity (Sodhi *et al.* 2004). For Borneo, currently only 57% of its land surface remains forested, and annual deforestation averages 1.7% (FAO 2006, Langner *et al.* 2007). Even more worrying is the fact that 56% of Kalimantan's (Indonesian Borneo) protected lowland forests has been lost between 1985 and 2001 (Curran *et al.* 2004, Stibig *et al.* 2007). Ca. 37% of Borneo's 15,000 vascular plant species (Roos *et al.* 2004) are thought to be endemic (van Welzen *et al.* 2005), with an estimated number of 10,000 species occurring in the WWF Borneo lowland rain forests ecoregion alone (Wikramanayake *et al.* 2002, Kier *et al.* 2005).

Considering the exceptional richness and concentration of endemic, or narrow ranged, species on Borneo, surprisingly little is known about the spatial distribution of both components. Only in 1995 the WWF and IUCN (1995) introduced the 'Centres of plant diversity' (CPD) for Australasia. In this contribution they argued that on Borneo most endemic plant species can be found in smaller areas in the north, the central mountain chain, and in the south-eastern Meratus Mountains (Fig. 4.1). A view largely supported by MacKinnon *et al.* (1996). Wong (1998) added to this list the 'Riau Pocket', which consists of two areas. One of these is similar to the north-western Sarawak biogeographical unit of MacKinnon *et al.* (1996), the other is the most western tip of Borneo (Fig. 4.1). Wong (1998) further suggests that Mount Kinabalu is a hotspot of plant diversity (Fig. 4.1), which is confirmed by its ca. 5,000 documented vascular plant species (Beaman 2005, Grytnes and Beaman 2006). Furthermore, Wong (1998) reports a comparatively lower diversity in the remaining area of Borneo, mainly consisting of the Kalimantan provinces.

These findings are partly confirmed by the only quantitative Borneo-wide study of lowland

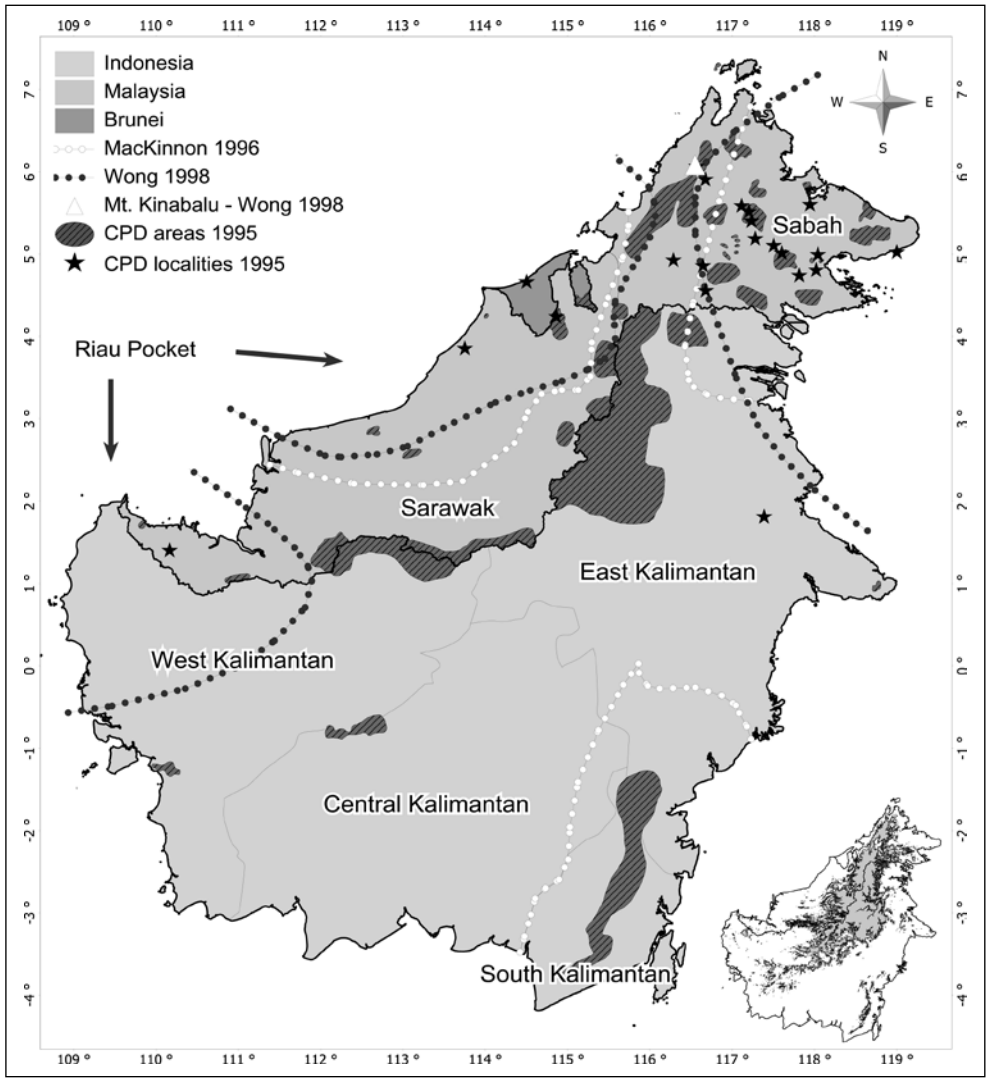


Figure 4.1. The country boundaries, the Indonesian and Malaysian provinces, the location of the centres of plant diversity (CPD) (WWF & IUCN, 1995), and the areas of high plant richness of Borneo by MacKinnon et al. (1996), and by Wong (1998). Inset shows all areas with an altitude of >500m asl.

dipterocarp forest (Slik *et al.* 2003). Based on data of 28 plots, at genus level, and for trees with a diameter of ≥ 10 cm, Slik *et al.*'s (2003) results only confirmed the biodiversity hotspots of the south-eastern Meratus Mountains and the north-western Sarawak biogeographical units. Their analysis did not support the comparatively lower diversity in the Kalimantan provinces of Wong (1998), however.
Except for the flora of Mount Kinabalu, and the plot studies by Slik *et al.* (2003), all the identified areas of high species richness and endemicity have boundaries that are largely based on informal expert opinions. Considering the rapid loss, and the lack of knowledge on the Borneo-wide spatial distribution of species richness and endemicity, warrants further studies.
With the ongoing digitization of natural history museum collections and herbarium specimens (Graham *et al.* 2004), much data has become available that allows for a quantitative analysis of species richness and endemicity patterns. Despite all collecting efforts of the last centuries, however, there is no complete inventory of all organisms inhabiting any single locality (Hortal *et al.* 2004). Moreover, collection localities are often biased towards easily accessible areas (Reddy and Davalos 2003, Kadmon *et al.* 2004, Hortal *et al.* 2007). Hence, the collecting effort on Borneo (and elsewhere) has been biased, as is evident from the spatial distribution of Bornean collection records (Raes and ter Steege 2007; Chapter 3, Fig.3.3). To overcome incomplete and biased sampling, and to be able to predict the geographic distributions of species even for areas where no collections have been made, has contributed to the development of species distribution modelling techniques (Araújo and Guisan 2006, Elith *et al.* 2006, Peterson 2006). A species distribution model (SDM) predicts the potential distribution of a species by interpolating identified relationships between presence/absence, or presence-only data of a species on one hand, and environmental predictors on the other hand, across an area of interest (Elith *et al.* 2006). Only few studies have attempted to develop species richness and endemicity patterns derived from SDMs (Zaniewski *et al.* 2002, Schmidt *et al.* 2005, Küper *et al.* 2006, Costa *et al.* 2007). Although these studies clearly contribute to the identification of possible centres of high richness, endemicity, and data deficiency; none of the underlying SDMs were statistically tested.
The importance of assessing whether an SDM differs from what would be expected on the basis of chance alone was first recognized by Olden *et al.* (2002). Recently, Raes and ter Steege (2007; Chapter 3) developed a methodology to test the significance of SDMs developed with presence-only data, the principal data of herbaria and natural history museums (Graham *et al.* 2004). Their methodology tests whether the correlations found between species' presence localities and the environmental predictors deviate from random chance expectation with a null-model (Raes and ter Steege 2007; Chapter 3).
To contribute to the conservation of botanical diversity of Borneo, we set out to model the patterns of botanical richness and endemicity, based on all significant SDMs at 5 arc-minute (~10x10 km at the equator) spatial resolution for all species treated in Flora Malesiana (Anon. 1959-2007) occurring on Borneo. Then, based on these patterns, we identify areas with higher levels of endemicity than can be expected on the basis of species richness. Finally, we analyse which environmental factors best explain the botanical richness and endemicity patterns.

Methods

Species data
We extracted all georeferenced species records from Borneo belonging to families treated in Flora Malesiana (Anon. 1959-2007) from the BRAHMS database of the National Herbarium of the Netherlands. To this dataset we added the georeferenced records of revised genera of the Annonaceae, Euphorbiaceae, and Orchidaceae. In total this dataset comprised 66,262 georeferenced records belonging to 102 plant families representing 4674 species. From this set of georeferenced records, we recorded

species presences for each 5 arc-minute grid cell, avoiding duplicate species records in one grid cell. We used a 5 arc-minute spatial resolution because this is the available resolution of the FAO soil property predictors (see below), and because georeferencing at a higher spatial resolution is not realistic. Furthermore, a species had to be represented in at least five grid cells to be modelled. These requirements were met for 2273 species represented by 44,106 unique records, ranging from 5 to 202 unique records per species.

Environmental predictors

To model the species distributions, we initially selected 37 environmental predictors. We downloaded the digital elevation model (DEM (in m)) and the 19 bioclimatic predictors (1950-2000) of the WORLDCLIM dataset (<http://www.worldclim.org>) for Borneo at 5 arc-minute resolution (Hijmans et al. 2005). Additionally, we selected 15 soil predictors from the FAO database for poverty and insecurity mapping (FAO 2002), shown in table 4.1. To this dataset we added a layer with the elevation range, defined as the difference between the lowest and highest altitude within a 5 arc-minute grid cell based on the 90m SRTM altitude data (<http://srtm.csi.cgiar.org>). Finally, a data-layer, reflecting the El Niño Southern Oscillation (ENSO) event drought impact, was added. ENSO drought impact was defined as the relative average annual difference in 'normalized difference vegetation index' (NDVI) values (<ftp://ftp.glcf.umiacs.umd.edu/glcf/GIMMS/Geographic/>) between months of a severe ENSO (07/1982-06/1983), and a non-ENSO year (07/1981-06/1982). These NDVI data were the oldest data available, and are therefore least affected by deforestation and land use change. We retained only grid cells with data for all data-layers, resulting in 8577 grid cells for Borneo. Records on the coast-line just falling outside the grid cells due to the 5 arc-minute resolution were shifted to their closest grid cell. Data-layer manipulations were performed with Manifold GIS (Manifold Net Ltd).

Table 4.1. Pearson's correlation coefficients (r) for the 15 FAO soil predictors and the five PCA soil axes. Values in italic r>0.4 or r<-0.4 (modest correlation) and in bold italic r>0.7 or r<-0.7 (strong correlation). Correlations are significant at 0.05 level (*), 0.01 Level (**), at 0.001 level (***), or not significant at 0.05 level (ns) (2-tailed). CEC = Cation Exchange Capacity; C:N = Carbon:Nitrogen.

	PCA01	PCA02	PCA03	PCA04	PCA05
Base saturation % topsoil	-0.1013 ns	***-0.8429*** ***	0.0974 ns	0.2864 ns	0.2268 ns
CEC clay topsoil	*0.5712* ***	-0.3342 *	*-0.5284* ***	0.4161 *	-0.1366 ns
CEC soil topsoil	***0.7449*** ***	-0.1708 ns	0.2366 ns	0.0295 ns	0.0806 ns
C:N-ratio class topsoil	*0.5083* ***	0.3100 ns	*-0.4183* **	-0.2982 ns	*0.5314* ***
Easy available water	***-0.7886*** ***	*0.4747* **	0.0332 ns	-0.1926 ns	0.1242 ns
Effective soil depth	0.2428 ns	-0.3322 *	-0.1498 ns	***-0.8224*** ***	-0.2733 ns
Nitrogen % topsoil	***0.7360*** ***	0.2317 ns	0.3555 *	0.1245 ns	-0.1529 ns
Organic carbon % topsoil	*0.5523* ***	*0.5221* ***	0.3205 *	0.0227 ns	-0.2646 ns
Organic carbon pool	***0.7626*** ***	0.3883 *	0.2412 ns	-0.1427 ns	0.1172 ns
pH topsoil	*-0.4389* **	*-0.6870* ***	0.1403 ns	0.2410 ns	-0.0953 ns
Soil drainage class	***0.8323*** ***	-0.2111 ns	0.1628 ns	0.1071 ns	0.2241 ns
Soil moisture storage capacity	***-0.7108*** ***	*0.5545* ***	-0.0116 ns	-0.0222 ns	0.1731 ns
Soil production Index	-0.0489 ns	***-0.8584*** ***	0.0836 ns	-0.1758 ns	-0.2444 ns
Textural class subsoil	-0.2891 ns	-0.0161 ns	***0.8747*** ***	-0.1090 ns	0.1550 ns
Textural class topsoil	-0.0382 ns	-0.1762 ns	***0.9153*** ***	-0.1784 ns	0.1307 ns

To avoid problems such as multi-collinearity (Graham 2003), which can result in model over-fitting (Peterson *et al.* 2007), we reduced the number of environmental predictors. From the bioclimatic predictors, together with the DEM, elevation range, and ENSO, we selected only those predictors which were least correlated (highest Pearson's r = 0.737; Table S4.1): DEM (correlated with and proxy for elevation range, and strongly negatively correlated with the mean annual temperature, maximum temperature warmest month and -quarter, minimum temperature coldest month and -quarter, mean temperature wettest- and driest quarter); Bio04 – Temperature seasonality; Bio07– Temperature annual range (correlated with diurnal temperature range); Bio12 – Annual precipitation (correlated with and proxy for precipitation in the wettest month and -quarter, driest month and - quarter, warmest - and coldest quarter); Bio15 – Precipitation seasonality; and ENSO (table S4.1). To reduce the number of soil predictors we used a principal component analysis (PCA). We performed the PCA on the 41 unique combinations of the 15 soil predictors values observed for the 8577 grid cells of Borneo. We selected the first five PCA-axes as our soil property predictors (PCA01-05), which together explained 83 % of the variance in the soil data. Pearson's correlation was used to determine which of the original 15 FAO soil predictors were significantly correlated to each of the five PCA axes (Table 4.1). This resulted in a reduction from 37 to 11 uncorrelated predictors, which were used to model the species distributions (Table S4.1; Fig. S4.1).

Species distribution model (SDM) building and testing with a bias corrected null-model

To model the species distributions we selected the modelling application Maxent (ver. 3.0.4; <http://www.cs.princeton.edu/~shapire/maxent/>) (Phillips et al. 2006). Maxent was specifically developed to model species distributions with presence-only data, has shown to outperform most other modelling applications (Elith *et al.* 2006, Pearson *et al.* 2007), is least affected by location errors in occurrences (Graham *et al.* 2008), and best performs when few presence records are available (Wisz *et al.* 2008). Maxent was set to use all species presence records for model building (explained below), by setting the 'random test percentage' to zero. The modelling rules were set to use linear features, when less than 10 records were available, adding quadratic features for SDMs developed with 10-14 records, and including hinge features for species with 15, or more, records (Raes and ter Steege 2007; Chapter 3). For each of the 2273 species an SDM was developed based on its unique presence records and the 11 environmental predictors.

As measure of SDM accuracy we used the threshold independent and prevalence insensitive area under the curve (AUC) of the receiver operating characteristic (ROC) plot (Fielding and Bell 1997, McPherson *et al.* 2004, Raes and ter Steege 2007; Chapter 3), produced by Maxent. All measures of SDM accuracy require absences (Fielding and Bell 1997). When these are lacking, as is the case here, they are replaced by pseudo-absences or sites randomly selected at localities where no species presence was recorded (Ferrier *et al.* 2002, Phillips *et al.* 2006). However, when SDM accuracy measures are based on presence-only data and pseudo-absences, the standard measures of accuracy (e.g. the often used measure AUC>0.7) do not apply (Raes and ter Steege 2007; Chapter 3). Therefore, we used the method presented in Raes and ter Steege (2007; Chapter 3) to test the AUC value of an SDM developed with all presence records against a bias corrected null-model of AUC values expected by chance. The AUC value of an SDM developed with ***n*** records is tested against

the upper 95% one-sided confidence interval (C.I.) AUC value derived from the AUC values of 1000 x *n* randomly drawn and modelled points. The random points were drawn from cells where in the past collections were made, and hence were corrected for any geographical sampling bias. For Borneo this was the case for 1837 (21.4%) of the total of 8577 grid cells (Raes and ter Steege 2007; Chapter 3).

We developed null-distributions for 5-35 records (31 distributions), for 40-50 records with intervals of 5 records (3 distributions), for 60-100 records with intervals of 10 records (5 distributions), and from 150-250 with intervals of 50 records (3 distributions). For each of these distributions we assessed the upper 95% one-sided C.I. AUC value, by ranking the 1000 AUC values and selecting the 950th value. We developed three series of C.I. values dependent on the modelling rules used by Maxent; 5-9 (only linear), 10-14 (linear & quadratic), and ≥ 15 (linear, quadratic & hinge) records. We applied a curve-fit to each of the three series against which the AUC values for all 2273 SDMs were tested. For further analyses only the significant SDMs were retained. To assess whether species represented by few records were not proportionally more often rejected than species with many records, we plotted the relative species abundance values against the relative species ranks. Similar shaped curves indicate that the sample is representative.

Additionally, we tested whether the 1837 collection localities were biased in environmental predictor space. For each of the 11 predictors we divided predictor space into 10 equal-interval bins based on the ranges observed for Borneo (8577 grid cells) (Loiselle et al. 2008). Then we tested whether the frequency distributions represented by the 1837 collection localities differed from all 8577 grid cells using a Chi-square test.

Botanical richness and endemicity patterns

In order to develop patterns of botanical richness and endemicity, a threshold was set to convert the continuous Maxent SDM predictions, which range from 0 to 100, to discrete presence/absence values. Although species identifications, and georeferencing of the collection localities, were carried out with the greatest possible accuracy, we found it reasonable to assume that 10% of the records were either wrongly identified, or georeferenced. Therefore, for all SDMs represented by ≥ 10 records we used the fixed '10 percentile presence' threshold. This threshold uses the Maxent value of the 10 percentile species presence record to define all areas with a lower predicted Maxent value as absent, and with a higher value as present. For those species represented by 5-9 records we used either the 'sensitivity-specificity equality' or the 'sum maximization' threshold (Liu *et al.* 2005), dependent on which of the two corresponding omission rate values was closest to 10%.

Once the thresholds were set, the botanical richness pattern was developed by superimposing all significant SDMs. To develop the endemicity pattern we used the weighted endemism index (Crisp *et al.* 2001, Kier and Barthlott 2001, Küper *et al.* 2006, Slatyer *et al.* 2007). This index weighs species richness according to the range sizes of the species present, and is calculated by summing the inverse of the range sizes of the species present in each grid cell. A species with a range of 10 grid cells has a weight of 1/10 in every grid cell where it is present. We developed the endemicity pattern by summing the weights of all significant SDMs for all grid cells.

Data analyses

We first assessed whether the modelled species richness did not under-predict the actual number of species collected in corresponding cells. This was done by plotting the predicted number of species against the collected number of species for all 8577 grid cells. To plot the data at log-log scale, 'the collected number of species' was transformed to the logarithm of 'the collected number of species + 1'.

The hotspots of endemic species were identified by mapping the relative residuals of the species richness – weighted endemism relationship. The relationship was assessed with a curve-fit procedure, evaluating several polynomial functions. The residuals of this relation were divided by their predicted weighted endemism values, resulting in a measure of relative residual weighted endemism.

To study which environmental predictors best explained species richness, weighted endemism, and their relative residuals we used a technique known as variation partitioning (Legendre 2008). Variation partitioning is a technique that partitions the variation of a response variable between two sets of explanatory variables, here a set of environmental predictors and a set of spatial predictors. As spatial predictors we used the nine terms of the third order polynomial trend-surface regression equation of latitude and longitude (Borcard et al. 1992, Lobo and Martin-Piera 2002, Legendre 2008). To account for possible non-linear effects between the diversity patterns and the original set of 11 environmental predictors, we added the quadratic terms of those predictors uncorrelated with the original predictors ($r<0.7$). This was the case for PCA02, PCA04 and PCA05 (Table S4.1). The variation was partitioned by performing a forward-backward stepwise multiple regression analysis for the three diversity measures and (1) the environmental predictors, (2) the spatial predictors, and (3) a combined matrix of both predictor sets. This allowed us to assess which proportion of the variation was attributed to only environmental predictors (a), to spatially structured environmental predictors (b), to spatial predictors (c), and to unexplained (residual) variation (d) (See Legendre 2008 for methodological details).

Multiple regression analyses applied to macroecological data which are driven by structured biological processes, may result in residual spatial autocorrelation (RSA). The presence of RSA is a violation of the assumption that residuals should be independent and identically distributed, and results in inflated Type I errors (Dormann et al. 2007). It was recently shown, however, that short-distance RSA, while causing inflated type I errors, does not seriously affect the interpretation of the regression coefficients estimated by ordinary least squares regressions (Diniz-Filho *et al.* 2007, Hawkins *et al.* 2007). To establish whether RSA was present in our regression residuals we assessed the Moran's I values with SAM – Spatial Analysis in Macroecology <http://www.ecoevol.ufg.br/sam> (Rangel *et al.* 2006). Since only a small proportion of the variation was explained by the spatial predictors alone (see Results), we estimated the regression coefficients for the three biodiversity measures based on the significant environmental predictors only.

Model performances were tested with a 10-fold cross-validation procedure, using 90% to train the model with the significant environmental predictors selected by the full-model, and test the model with the remaining 10%, repeated 10 times. All regression analyses were performed using SPSS version 15 (SPSS Inc.).

It can be argued that this approach has a certain degree of circularity in reasoning, since the same predictors were used to develop the underlying SDMs. We argue, however, that Maxent identifies correlations with the environmental predictors independently for each species. Even if for two species the same

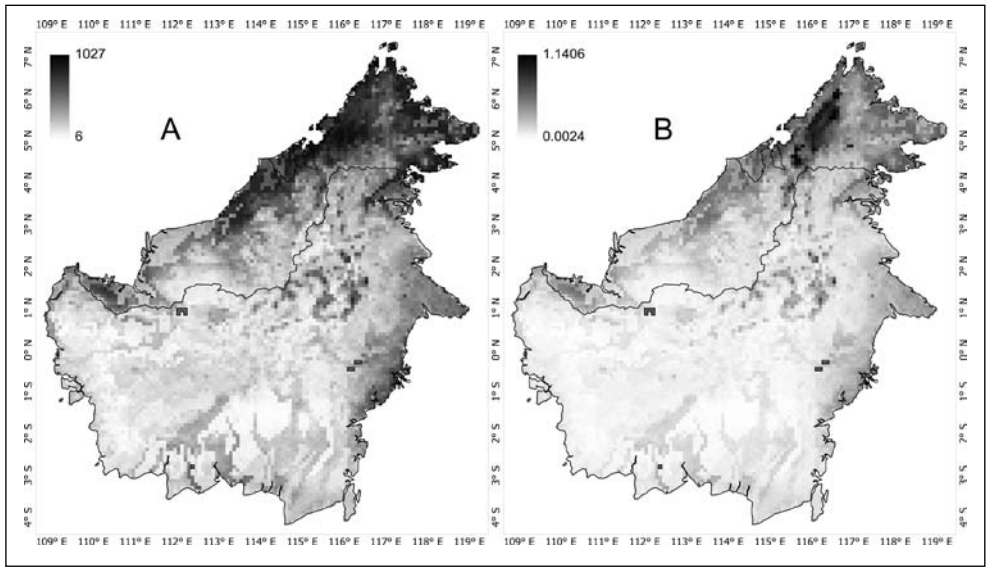


Figure 4.2. Botanical richness (A) and weighted endemism (B) pattern derived from 1439 superimposed significant species distribution models.

predictors are selected as being the most important to predict their distributions, these two species can have another optimum in their response. So we posit that the significant SDMs underlying the superimposed species richness, weighted endemism, and relative residual weighted endemism patterns are essentially independent.

Results

From the 2273 species which were modelled, 1439 (63.3 %) had a distribution pattern that differed significantly from a random one. The frequency distributions of the environmental conditions represented by the collection localities did not significantly differ from the distributions for all 8577 grid cells of Borneo (Fig. S4.2). The relative species rank abundance curves (Fig. S4.3) for the records of all 2273 species, and for the records of the 1439 significant species, were largely similar in shape. This indicated that in terms of the frequency distributions of records, the 1439 species represented by a significant SDM are a representative sample of the total of 2273 modelled species.

Superimposing the significant SDMs, and derived weighted endemism values, resulted in the botanical richness and weighted endemism maps as presented in Fig. 4.2a and b, respectively. The highest richness was found in Sabah, where 1027 species of the total of 1439 species were predicted to occur in one grid cell. The lowest predicted richness was found in West Kalimantan where only 6 of the modelled species were predicted to be present. Besides Sabah and north-western Sarawak, also the high mountains in East Kalimantan had very high predicted richness values (Fig. 4.2a). The weighted endemism values (Fig. 4.2b) showed largely the same pattern as the richness pattern, albeit more concentrated around the mountains, notably the northern Crocker Range, and not extending into the lowland.

The log-log plot of the predicted number of species plotted against the species counts per grid cell, showed that for cells where few species were actually collected, almost the whole range of predicted numbers of species was found (Fig. 4.3). For cells with highest numbers of actually collected species, however, the predicted number of species was never lower. Only for 2 grid cells (0.13%) a lower number of species was predicted than were actually collected (Fig. 4.3; below diagonal and inset).

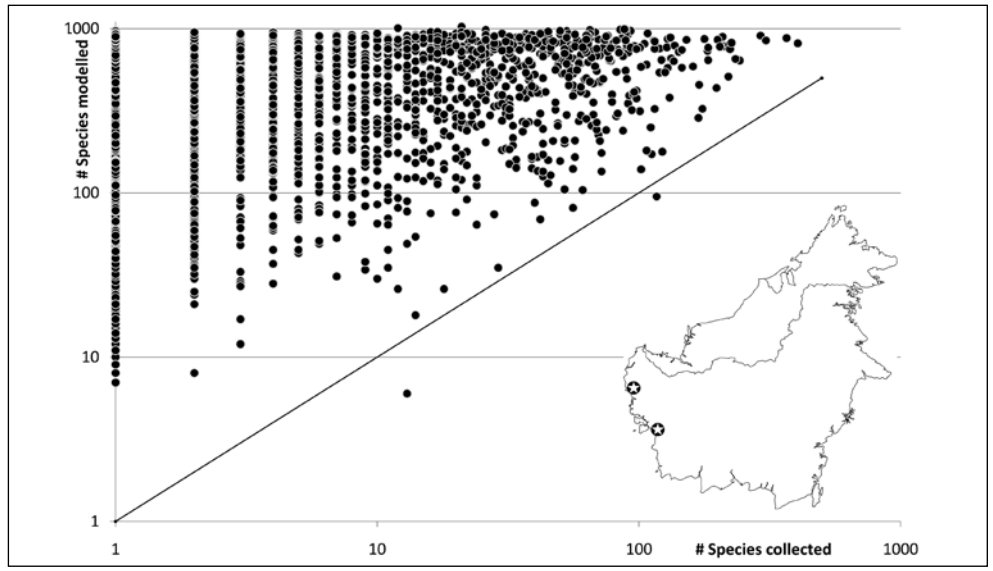


Figure 4.3. The predicted number of species plotted against the number of species actually collected per grid cell based on species which had a significant species distribution model. The two grid cells for which a lower number of species was predicted than actually was found (below the diagonal line) are geographically plotted on the inset. Note the log scale of both axes.

The relationship between species richness and weighted endemism was best described by a quadratic function ($R^2 = 0.914$; $p<0.001$; Fig. 4.4 bottom). Mapping the relative residuals of this relationship, exposed those areas with higher weighted endemism values than can be expected on the basis of species richness (Fig. 4.4 top; black and white areas). These areas were located on Mount Kinabalu and the Crocker Range Mountains in the north, the Müller Mountains in the south of the central mountain chain, on the east-side of the southern Meratus Mountains, and in the north of East Kalimantan's Sangkulirang peninsula.

The results of the variation partitioning showed that the majority of the variation in species richness and weighted endemism can be explained by spatially structured environmental variables (Fig. 4.5; fraction b). Only very small fractions (3.4-4.5%) were attributed to spatial variables alone (Fig 4.5; fraction c). Hence, the regressions were performed on the environmental predictors only. The forward-backward stepwise regression results for both species richness and weighted endemism suggest that the most important environmental variables were Bio07 – temperature annual range, Bio04 – temperature seasonality, and ENSO (Table 4.2). Together all significant environmental variables explained 82.8, and 75.3 percent of the total variance in species richness and weighted endemism, respectively (Table 4.2; Fig. 4.5, fraction a+b). For the

Table 4.2. Results of the forward-backward stepwise multiple regression for species richness, weighted endemism, and the relative residual weighted endemism values as a function of the environmental predictors (n=8577). The order indicates the sequence in which the significant environmental predictors (Pred.; p<0.05) were included in the regression equation. $R^2_{adj.}$ indicates the cumulative adjusted coefficients of determination after the inclusion of each variable. Beta is the standardized regression coefficient, and t is the corresponding t-value for the full regression model. All models are significant at p<0.001. Legend: DEM – Digital Elevation Model; Bio04 – Temperature seasonality; Bio07 – Temperature annual range; Bio12 – Annual precipitation; Bio15 – Precipitation seasonality; PCA01-05 – PCA soil axes 1-5; ENSO – Drought impact by El Nino- events on NDVI values.

Species Richness				Weighted Endemism				Residual Weighted Endemism			
Pred.	R^2_{adj}	Beta	t	Pred.	R^2_{adj}	Beta	t	Pred.	R^2_{adj}	Beta	t
Bio07	0.123	-0.931	-132.034	**Bio07**	0.107	-0.840	-96.536	**DEM**	0.236	0.647	65.472
Bio04	0.546	1.026	134.268	**Bio04**	0.456	0.975	104.508	**PCA05²**	0.352	0.311	38.693
ENSO	0.640	0.258	49.341	**ENSO**	0.547	0.236	36.006	**PCA03**	0.409	-0.182	-23.512
PCA05²	0.695	-0.211	-41.642	**DEM**	0.599	0.345	45.455	**Bio12**	0.460	-0.241	-28.976
PCA03	0.745	0.223	42.251	**PCA05²**	0.635	-0.171	-27.745	**PCA04**	0.486	-0.219	-27.414
PCA05	0.774	0.225	47.701	**Bio12**	0.674	-0.240	-23.337	**Bio15**	0.518	0.195	20.319
PCA02²	0.787	-0.151	-28.427	**PCA05**	0.700	0.194	33.973	**PCA01**	0.535	0.155	19.263
Bio12	0.807	-0.157	-30.488	**PCA03**	0.721	0.158	24.581	**ENSO**	0.544	0.104	12.026
DEM	0.822	0.171	27.076	**PCA02²**	0.747	-0.156	-23.337	**PCA05**	0.551	-0.105	-13.893
PCA02	0.826	-0.044	-7.674	**PCA02**	0.750	-0.037	-5.420	**PCA02²**	0.557	0.087	9.890
PCA04²	0.827	-0.053	-8.578	**PCA04**	0.751	-0.022	-3.350	**Bio07**	0.561	0.149	13.011
PCA01	0.828	-0.036	-6.854	**PCA01**	0.751	-0.036	-5.696	**Bio04**	0.565	-0.105	-8.528
				PCA04²	0.752	-0.042	-5.163	**PCA02**	0.566	-0.040	-4.917
				Bio15	0.753	0.028	3.800				

residuals, DEM (altitude), the quadratic term of PCA soil axis 5, and PCA soil axis 3 were the most important environmental variables. In total the model explained 56.6% of the variance in relative residual endemism (Table 4.2).
The 10-fold cross-validation results of the models obtained through stepwise regression showed that the average $R^2_{adj.}$ values of the test-data partitions were not significantly different from the average training-data partition values (Fig. 4.6), implying that the models were not over-parameterized. The Moran's I values of the regression residuals of species richness and weighted endemism indicated that some RSA was still present for the first three lags (Fig. 4.7).

Discussion

Botanical richness pattern

The richness pattern is based on 1439 significant SDMs, approximately 10% of the estimated number of 15,000 species expected to occur on Borneo. This is the largest dataset available today and represents all life-forms from herbs to canopy trees represented by 102 plant families treated in Flora Malesiana. The relative species rank abundance curve of the species represented by the 1439 significant SDMs lies slightly above the curve for all 2273 species that were modelled (Fig. S4.3). This indicates that a smaller proportion of species represented by few records is represented in the diversity pattern. It was shown, however,

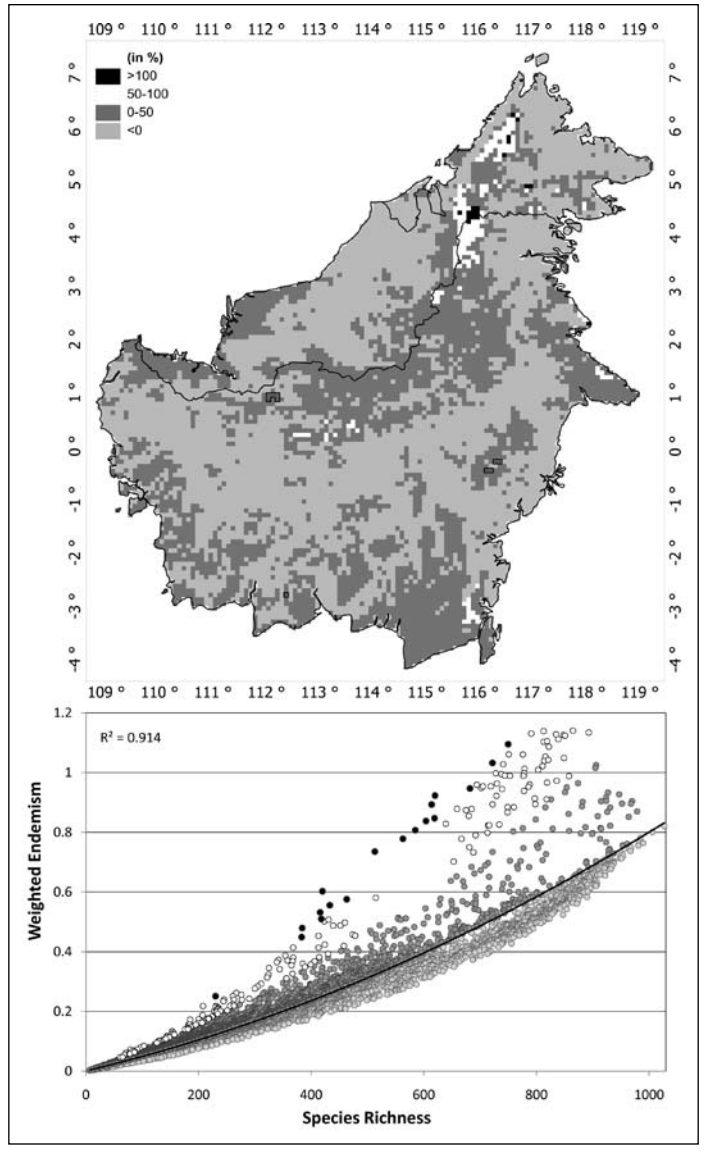


Figure 4.4. The relationship between botanical richness and weighted endemism. Light grey dots represent grid cells with negative relative residual weighted endemism values; dark grey dots positive relative residual endemism values between 0-50%, white dots between 50-100%, and black dots >100%. Residual classes are mapped in the top image.

that commoner species are most responsible for richness patterns (Lennon *et al.* 2004). We argue therefore that this is not problematic for the interpretation of the botanical richness pattern. The implications for the endemicity patterns are discussed below.
The highest botanical richness is predicted to occur in northern, and north-western Borneo (Fig. 4.2a). These results confirm the suggestion that northern Borneo is a centre of vascular plant diversity (Barthlott *et al.* 2005, Kier *et al.* 2005, Mutke and Barthlott 2005). In the lowlands the highest species richness was found in the areas west of the northern Crocker

Mountains range, while the rest of the lowlands had much lower predicted richness values. This distribution was previously suggested by Wong (1998). Furthermore, we expect high richness values for the transition area between the Sembakung-Sesajap delta and the lowland forest of northern East Kalimantan, and for the most western tip of Sarawak, around Kuching. All these areas fall within the roughly demarcated areas of high richness by Mackinnon *et al.* (1996) and Wong (1998) (Fig. 4.1). Additionally, our modelled richness pattern identified the eastern mountain peaks of the central mountain chain located in central East Kalimantan (Fig. 4.2a) as an area of potentially high richness, previously not recognized. Our results cannot support the south-eastern Meratus Mountains as an area of exceptional plant species richness as was indicated by MacKinnon *et al.* (1996).
The lowest plant species richness values were found in the southern, and south-western parts of Borneo (Fig. 4.2a). These areas largely coincide with the WWF Sundaland heath- and peat swamp forest ecoregions (Wikramanayake *et al.* 2002). According to Wikramanayake *et al.* (2002) these forest types are generally less species-rich than comparable lowland dipterocarp forest. Similarly, the peat swamp forests along the coast of Sarawak, and part of Brunei, also have lower predicted species richness values than the surrounding lowland forest. Nevertheless, we expect that the predicted richness for the lowland regions of the Kalimantan provinces still underestimates the diversity due to relatively low collection densities in this region (Raes and ter Steege 2007; Chapter 3, Fig.3.3). For a species to be modelled it should be represented by at least 5 unique records in our data set. The probability that Kalimantan endemics are represented by 5 records is much lower than for north-western Borneo endemics. On the other hand, the 10 equal-interval bins frequency distributions of the environmental predictors for the collection localities did not differ from those of all 8577 grid cells (Fig. S4.2). A lower predicted number of species caused by underrepresentation of environmental conditions represented by the collection localities is therefore not expected. At the same time South-, East-, and Central Kalimantan represent regions most affected by the ENSO (Slik 2004) (Fig. S4.1), consistent with the lower richness values predicted (explained below).
One of the two localities that had a lower predicted number of species (95) than were actually found (117) is located in the Gunung Palung National Park (1°15´S, 110°10´E; Fig.4.3 (inset)). The 15 km^2 study site in this park harbours seven distinctive forest types (Cannon *et al.* 2007). The resolution of our analysis is 100 km^2, and therefore cannot correctly represent the different environmental conditions present in the park.

Weighted endemism pattern and centres of endemicity

The distribution of narrowly ranged-, or endemic species is often regarded as a metric of higher importance for conservation planning than species richness (Reid 1998, Myers *et al.* 2000). The fact that a smaller proportion of species represented by few collections is represented by the significant SDMs (Fig. S4.3), implicates that endemicity values are expected to be even higher than presented in this study. The northern Crocker Mountains range with Mount Kinabalu, and the high mountains of central East Kalimantan have the highest weighted endemism values (Fig. 4.2b). The latter have received little collecting effort so far, and deserve further attention since they potentially harbour many new species.
Similar to our data (Fig. 4.4), a curvilinear relationship between richness and endemicity was also found for the birds of Africa (Jetz *et al.* 2004), indicating that more endemic species can be expected in species rich assemblages (Witt and Maliakal-Witt 2007). Spatial mapping of the relative residuals of this relationship revealed the centres of endemicity on Borneo (Fig. 4.4): the Crocker Mountains range with Mount Kinabalu; the northern parts of the central mountain range; the high mountain peaks in east Sabah; the southern extrusions of the central mountain range (Müller Mountains); the lowland east of the southern Meratus Mountains; and the eastern Sangkulirang peninsula of East Kalimantan. It is notable that our results add the Müller Mountains and the Sangkulirang peninsula to the previously known list of biologically important sites on Borneo (Fig 4.1).
Besides the entire central mountain range having positive residual weighted endemism values, this is also the case for south-west Sarawak, the southern, and south-western areas of Borneo, and around the great lakes in southern East Kalimantan. Although the absolute richness and weighted endemism values for these areas are low, they apparently harbour species which are very characteristic for those areas, and are not found elsewhere (Fig.4.4). These areas largely coincide with the WWF peat swamp-, freshwater swamp-, and heath forest ecoregions of Borneo (Wikramanayake *et al.* 2002).

Botanical richness, weighted endemism, and centres of endemicity explained

The results of the variation partitioning showed that for all three diversity measures only a very small proportion of the variance is explained by spatial predictors only, and that for species richness and weighted endemism the majority is explained by spatially structured environmental variables (Fig. 4.5). Although the Moran's I values of the residuals from the partial regressions with only environmental variables were slightly higher than for the models including only spatial-, or spatial and environmental variables combined (Fig. 4.7), they do fall well within the ranges reported by Hawkins *et al.* (2007). They concluded that for these ranges of RSA regression coefficients were not seriously affected. Therefore we analysed the diversity patterns with the environmental predictor dataset only (Table S4.1; Fig. S4.1). The 10-fold cross validation results suggest that the predictive models for the three diversity measure performed well and were not over-parameterized, since none of the average test $R^2_{adj.}$ values differed from the average training values (Fig. 4.6).
The most important variable, when tested alone, explaining most of the richness pattern

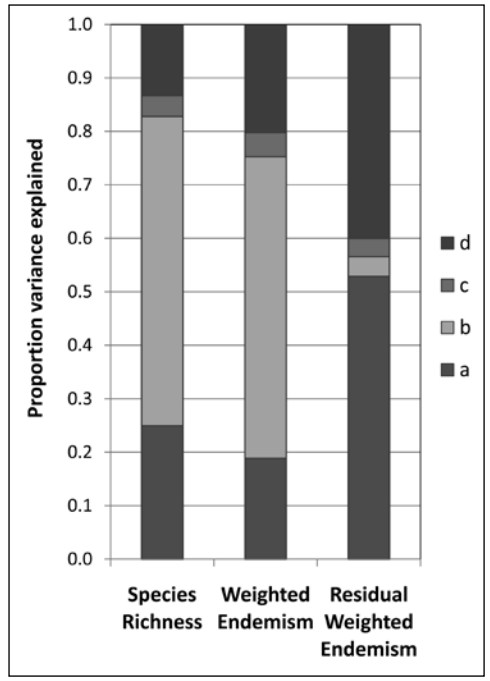


Figure 4.5. The results of variation partitioning of the forward-backward stepwise multiple regressions for species richness, weighted endemism, and relative residual weighted endemism values; a = variance explained by environmental factors, b = variance explained by spatially structured environmental factors, c = variance explained by spatial factors, and d = unexplained (residual) variance.

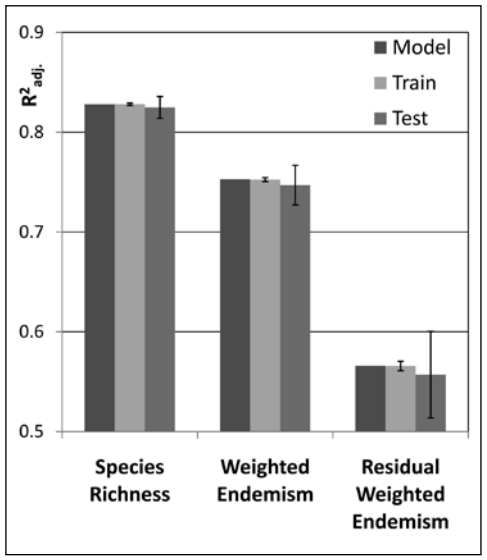


Figure 4.6. The adjusted R^2 values of the 10-fold cross-validation results for species richness, weighted endemism, and the relative residual weighted endemism values. Error-bars indicate ± 2SD. Average train and test $R^2_{adj.}$ values were not significantly different (two-sample t-tests assuming unequal variances).

was the annual temperature range (Table 4.2; Bio07). The negative correlation with this variable suggests that the highest richness values were found under relative stable annual temperature conditions. The variable explaining most of the variance in species richness was temperature seasonality (Table 4.2; Bio04). This variable was positively correlated with species richness, which suggests that seasonality in temperature may be a driving factor of species richness. It should be noted however, that temperature seasonality, expressed as the standard deviation of weekly mean temperatures as a percentage of the mean annual temperatures, only ranged from 1.11-5.37%. The same two variables also accounted for almost 50% of the total explained variance of the weighted endemism pattern (Table 4.2). Stable climatic conditions maintaining high richness and endemicity values have been found for various organisms on different continents. For the birds of Africa, low seasonality, best captured through the annual temperature range, was found to be the second most important predictor for centres of endemism (Jetz *et al.* 2004). For Amazonia, the highest botanical richness was found in areas with the shortest dry season length (ter Steege *et al.* 2003). It can be argued that habitats which face a long dry season have a larger difference in temperature between dry and wet months than habitats which remain wet throughout the year. For reptiles and amphibians in Europe, both temperature and precipitation stability were found to be important predictors of high species richness (Araújo *et al.* 2008). Araújo *et al.* (2008) even showed that it is not only contemporary climatic stability which maintains high species richness, but that stability in climate since the last glacial maximum (LGM) is an even better predictor. Similar results were found for the Australian wet tropics, predicting the highest number of species for a number of taxonomic vertebrate groups in areas which have remained climatically stable since the LGM (Graham *et al.* 2006). For Borneo there are only indirect suggestions that the areas of high richness and endemicity have been stable in temperature and precipitation over longer time-scales. Geomorphic evidence suggests drier, cooler, and more seasonal climates during the LGM (Verstappen 1997), which resulted in a savanna corridor running from the southern, and south-western areas of Borneo, through the present-day Java sea and Karimata street, into south-east Asian mainland during that period (Heaney 1991, Gathorne-Hardy *et al.* 2002, Bird *et al.* 2005). There are strong indications, however, that northern Sarawak, Brunei, Sabah and East Kalimantan up to the Barito river remained forested, with everwet conditions in northern Borneo and lowland rainforest surviving around montane rainforest patches (Gathorne-Hardy *et al.* 2002). These are the areas which coincide with the areas where the highest richness and endemicity values are predicted today.

The mechanism by which temperature seasonality (Bio04) drives high species richness and endemicity values remains speculative. There is a possible relation to phenological diversity driven by seasonal differences in abiotic conditions such as temperature and humidity (Sakai 2001). Temporal segregation of flowering minimizes interspecific overlap in flowering times, and thus ineffective pollination, or competition for pollinators. This hypothesis in known as the shared pollinator hypothesis (Sakai 2001). Whether seasonality in temperature (Bio04) within areas with a small annual temperature range (Bio07) has a clear seasonal pattern remains to be investigated, however.

Another factor of importance, explaining 9.4% and 9.1% of the variance in species richness and weighted endemism, respectively, was the ENSO drought predictor (Table 4.2; Fig. S4.1). The highest richness values were found in areas least affected by ENSO, which could indicate that severe ENSO drought impact leads to local extinction. This could also explain why the richness values for the Kalimantan lowland areas are lower than for those in north-, and north-west Borneo. These findings are supported by plot studies in East Kalimantan that found disproportionate mortality of certain species groups and tree size classes during the severe ENSO event of 1997/1998 (Slik 2004). The lower species richness values in the southern, and south-western areas of Borneo identified by the WWF as Sundaland heath- and peat swamp forest ecoregions (Wikramanayake *et al.* 2002), are explained by variables PCA052 and PCA03 (Table 4.2). Heath forests, or kerangas, are commonly found on soils known as white-sand soils, and are often covered by a layer of peat or humus. Peat swamp forests form when sediments and organic matter builds up behind mangroves. The peat deposits can extend up to 20m (Wikramanayake *et al.* 2002). Besides along the southern coast, peat swamp habitats are also found in west Sarawak, Brunei, and around the lakes in south East Kalimantan. The variable PCA05

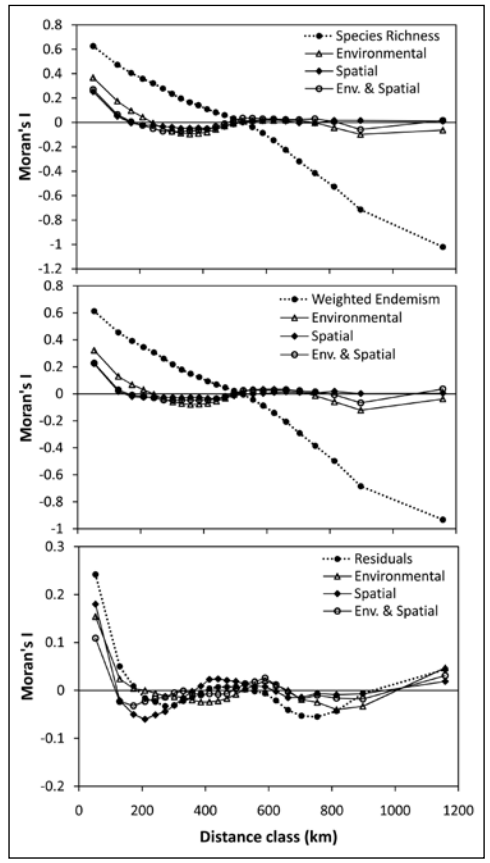


Figure 4.7. Moran's I values (equal number of pairs lags) for: species richness (A), weighted endemism (B), and the relative residuals weighted endemism values (C) (closed circles); and for the residuals of the forward-backward stepwise multiple regression results with only environmental predictors (open triangles), only spatial predictors (closed diamonds), and environmental and spatial predictors combined (open circles).

was positively correlated with the C:N-ratio of the topsoil (Table 4.1). The negative relation of PCA052 to species richness indicates that intermediate carbon content over nitrogen, characteristic for peat swamps, may have a negative effect on species richness. PCA03 was positively correlated with the textural class of the top- and subsoil (Table 4.1). The identified areas have low values for both predictors, which corresponds with coarse-textured sandy soils (FAO 2002), characteristic for kerangas and peat swamps (Whitmore 1984). Poor soil conditions, relative isolation, and the likely presence of a savanna corridor during the LGM may have resulted in low present day richness values.

The factor accounting for most of the explained variance in relative residual endemism values is the DEM (Table 4.2; Fig.4.4). Amongst others, the altitudinal range is correlated with this variable (see Methods). A large altitudinal range was also identified as the most important variable explaining the centres of African (Jetz *et al.* 2004), and South-American bird endemism (Rahbek *et al.* 2007). Jetz *et al.* (2004) argued that topographic heterogeneity might be better viewed as "a rough surrogate variable reflecting historical opportunities for allopatric speciation", which can result in centres of endemism. The mechanism suggested to drive speciation is the occurrence of narrow homothermal elevation bands serving as past and present dispersal barriers (Jetz *et al.* 2004). Other variables explaining a substantial portion of the variance in residual endemism values were PCA052, PCA03, and annual precipitation (Bio12). Where PCA052 and PCA03 explained low species richness values, the signs of the relation of these variables to relative residual endemism values were inverse, suggesting that the corresponding conditions promote speciation, resulting in positive residual weighted endemism values for the heath- and peat swamp forests (Table 4.2). Although annual precipitation (Table 4.2; Bio12) only explained 5.1% of the variance, the annual precipitation pattern (Fig. S4.1; Bio12) showed large similarities with the pattern of the centres of endemicity (Fig. 4.4). High relative residual endemism values were found in areas with the lowest annual precipitation. All the areas are separated by wetter areas, effectively isolating the dryer areas, which might have promoted speciation.

With this study, we quantitatively analysed the Borneo wide, high-resolution botanical diversity and endemicity patterns. We showed that herbarium records can effectively be used to develop these patterns, covering areas that never have been botanically sampled. The analysis predicted an additional centre of high diversity and endemicity for the mountains in East Kalimantan, and two additional centres of endemicity; the southern extrusions of the central mountain chain, the Müller Mountains, and another on the Sangkulirang peninsula. Furthermore, our results quantitatively confirmed many of the previously recognized areas of high botanical richness and endemicity, which were based on informal expert opinions. The variables explaining most of the variance of the three biodiversity measures were comparable to other macroecological diversity studies, an indication for the reliability of our results. Additionally, our results suggested that the centres of endemicity were best explained by ecological isolation. The variables involved were altitude, soil types, and annual precipitation.

Although we are confident that the estimated patterns reflect the true richness and endemicity patterns, we also stress that areas with lower values for the three diversity measures are not necessary less important for conservation. These areas may harbour species not found elsewhere, or have a forest composition, which is different from the ones found in the 'hotspots'. We hope that our results will guide conservation efforts for the severely threatened forests of Borneo.

Acknowledgements - We like to thank Jaap Vermeulen for sharing his orchid database, the NWO-Groot 'Building databases for life'-project (Grant# 175.010.2003.010) for funding the digitization of all Borneo collections used for this research, and two anonymous reviewers for valuable comments on the manuscript.

CHAPTER 5

The floristic regions of Borneo inferred from species distribution models

Niels Raes, E. Emiel van Loon, Marco C. Roos and Hans ter Steege

Submitted

Abstract

Aim

A quantitative floristic analysis of Borneo using herbarium records to justify the recognition of floristic regions.

Location

Borneo; between approx. 4°S-7°N and 109-119°E.

Methods

From the collection of vascular plants of Borneo stored at the National Herbarium of the Netherlands, we extracted records (N=44,106) belonging to families revised in Flora Malesiana (including revised genera of the Annonaceae, Euphorbiaceae, and Orchidaceae). To limit the influence of collection bias, we modelled each species' distribution using the presence-only modelling technique 'Maxent' in combination with 11 environmental independent variables at 5 arc-minute resolution (ca. 100 km^2 grid cells). The significance of each species distribution model (SDM) was tested against a bias corrected null-model. We constructed a presence/absence matrix based on 1439 significant SDMs (63.3 % of the 2273 species tested) for the 8577 grid cells of Borneo. This matrix was then analysed using a hierarchical cluster analysis, and the resulting cluster dendrogram was pruned using indicator species analysis (ISA) to partition floristic regions. The relationship between floristic regions and environmental conditions was then explored using a classification and regression tree (CART) analysis.

Results

The cluster analysis identified 11 floristic regions for Borneo. Many of the regions overlap with previous classifications based on informal expert opinion. Our analysis indicates, however, that the lowland rain forest region can be divided into at least four distinct floristic sub-regions. The CART analysis identified meaningful ecological thresholds defining each floristic region, largely in accordance with the known ecology of each floristic region.

Main conclusions

Our collections-based analysis identified 11 floristic regions of Borneo. Results largely confirm the floristic distinctiveness and extent of montane rain forest, kerangas, peat swamp, and fresh water swamp forest. The lowland rain forest, previously recognized as one floristic region is divided in at least four (and possibly six) distinct regions, viz. the lowlands of (i) Sabah and Sarawak, (ii) East Kalimantan, (iii) southern Borneo, and (iv) the Wet hill forest of Sarawak. The 'Kinabalu highlands', mangroves, and forests on lime-stone and ultramafic rock could not be distinguished due to the 100 km^2 resolution of our analysis.

Keywords

Phytogeography; Borneo; species distribution model; Maxent; significance test; null-model; bias; cluster analysis; indicator species analysis; classification and regression trees.

Introduction

Borneo is one of the most important biodiversity hotspot of the world (Myers *et al.*, 2000; Ashton, 2005; Kier *et al.*, 2005), and is severely threatened by deforestation and land use change (Curran *et al.*, 2004; Langner *et al.*, 2007; Stibig *et al.*, 2007). Although this is widely recognized, Borneo's internal floristic division, defined as areas that are relatively homogeneous in plant species composition, is still largely based on informal expert opinion. With ongoing global efforts to digitize herbarium records, an increasing amount

of data on the spatial distribution of species' occurrences has become available (Graham *et al.*, 2004). This creates new opportunities to analyse quantitatively the floristic structure of biodiversity hotspots such as Borneo, and thereby support conservation efforts.

Between 6-8 floristic regions have previously been recognized for Borneo (van Steenis, 1958b; Whitmore, 1984b; MacKinnon, 1997; Wikramanayake *et al.*, 2002). The first attempt to define the floristic structure of Borneo was made by Van Steenis (1935). For the entire country of Indonesia he recognized nine floristic regions, six of which occurred on Borneo (van Steenis, 1935a). Since then, various additions and improvements to the floristic map of South-East Asia, including Borneo, have been made (Hannibal, 1950; van Steenis, 1958a, b; Whitmore, 1984b; MacKinnon, 1997) (Fig. S5.2). The most recent of these is the WWF ecoregion map of Borneo, described in detail by Wikramanayake *et al.* (2002) (Fig. S5.2), which recognizes seven ecoregions for Borneo, based on the general framework of MacKinnon's (1997) 'bio-units'. WWF departed from MacKinnon's bio-units in three ways: a) ecoregion delineations were more closely based on potential vegetation maps, whereas vertebrates were the dominant influence on MacKinnon's bio-units; b) mixing of natural communities that are characteristic of specific habitat types was avoided; and c) lowland forests were separated from montane forests (Wikramanayake *et al.*, 2002).

The emerging consensus from past efforts is that at least six distinct floristic regions can be distinguished on Borneo: montane rain forest, lowland evergreen rain forest, heath forest, freshwater swamp forest, peat swamp forest, and mangrove forest. Additionally, Whitmore (1984) and MacKinnon (1997) recognize limestone forest formations, which were not mapped by WWF due to uncertainties concerning location and geographical extent (Wikramanayake *et al.*, 2002). WWF also recognizes the 'Kinabalu montane alpine meadows' as an separate ecoregion (Wikramanayake *et al.*, 2002), and the ultramafic outcrops of eastern Sabah as a distinctive flora (WWF & IUCN, 1995). The only quantitative floristic analysis of lowland flora of Borneo based on data of 28 plots, recognized five major floristic lowland regions, i.e. Santubong/Bako cluster of western Sarawak, eastern Sabah, central Sarawak, the Kalimantan provinces, and a northern cluster covering northern Sarawak and Brunei with an extension into northern East Kalimantan (Slik *et al.*, 2003). These results illustrate the potential for quantitative analyses to uncover hitherto unseen patterns of geographic variation, even within ecoregions, but to date such attempts are limited to that of Slik *et al.* (2003).

The volume of digitized herbarium records now make it possible to undertake quantitative analyses of floristic variation across very large spatial scales (Linder *et al.*, 2005). Unfortunately, collection localities are generally not uniformly spatially distributed (Kadmon *et al.*, 2004; Moerman & Estabrook, 2006; Hortal *et al.*, 2007), as was shown for Borneo (Raes & ter Steege, 2007 - Chapter 3). Efforts to overcome problems of geographically biased collection localities in a data set to be used for predicting species presence/absence across areas where no collections have been made has lead to development of a suite of species distribution modelling techniques (Guisan & Zimmermann, 2000; Elith *et al.*, 2006; Peterson, 2006). Species' distribution models (SDMs) predict the potential distribution of species by describing relationships between species' presence/absence-, or presence-only data, and environmental predictors across an area of interest. If meaningful environmental predictors are available at the appropriate resolution, in combination with sufficiently accurate collection localities, SDMs can predict the presence and absence of species across the entire area of investigation at the spatial resolution of the environmental predictors.

Considering the threatened status of Bornean forests, the variety of past efforts and results to describe floristic divisions, and recent advances made in the development of SDM methods, we aimed to: a) construct species presence/absence maps at 5 arc-minute (ca. 100 km^2) resolution for a large number of plant species using SDMs; b) delimit quantitatively the floristic regions of Borneo, by analysing the complete set of significant species presence/absence maps using hierarchical cluster analysis; and c) identify the ecological factors that drive the formation of Borneo's floristic regions.

Materials and Methods

Species data and environmental predictors

We extracted all georeferenced species records from Borneo belonging to the families treated in Flora Malesiana (Anon., 1959-2007) from the BRAHMS database of the National Herbarium of the Netherlands. We added the georeferenced records of revised genera of the Annonaceae, Euphorbiaceae, and Orchidaceae. This dataset comprised 66,262 georeferenced records belonging to 102 plant families. Only species with records in five, or more, grid cells were included. The resulting data set included 2273 species representing 44,106 unique records, ranging from 5 to 202 records per species.

Initially, 37 environmental predictors were selected to model species distributions. We downloaded the digital elevation model (DEM (in m)) and the 19 bioclimatic predictors (~1950-2000) of the WORLDCLIM dataset (<http://www.worldclim.org>) for Borneo at 5 arc-minute (ca. 100 km^2) resolution (Hijmans *et al.*, 2005). Additionally, 15 soil property

Table 5.1. Pearson's correlation r values for the 11 environmental predictors used to model the species' distributions, and the uncorrelated quadratic terms additionally used in the CART. Legend: DEM - Altitude; Bio04 - Temperature seasonality; Bio07 - Temperature annual range; Bio12 - Annual precipitation; Bio15 - Precipitation seasonality; PCA01-05 - PCA soil axes 1-5; ENSO - El Nino drought impact.

	DEM	Bio04	Bio07	Bio12	Bio15	PCA01	PCA02	PCA03	PCA04	PCA05	ENSO	PCA02^2	PCA04^2
Bio04	-0.512												
Bio07	-0.361	**0.737**											
Bio12	0.207	-0.167	0.030										
Bio15	-0.412	0.514	0.507	-0.254									
PCA01	-0.216	0.194	0.142	-0.030	0.185								
PCA02	0.274	-0.078	0.094	0.214	0.040	0.138							
PCA03	0.226	-0.214	-0.202	0.066	-0.168	-0.306	-0.036						
PCA04	-0.064	0.026	0.081	-0.095	0.149	0.034	0.151	-0.018					
PCA05	-0.136	0.033	0.004	-0.062	0.048	0.068	-0.011	0.021	-0.010				
ENSO	0.372	-0.078	-0.099	0.274	-0.390	-0.053	-0.003	0.128	-0.083	0.034			
PCA02^2	0.078	-0.035	-0.023	-0.208	-0.026	-0.028	0.044	0.022	0.364	0.230	0.084		
PCA04^2	-0.085	0.134	0.193	-0.109	0.184	-0.230	0.323	0.285	0.464	0.119	-0.126	0.331	
PCA05^2	-0.062	0.149	0.112	-0.140	0.094	-0.170	0.060	-0.078	0.049	0.112	-0.058	0.334	0.292

Table 5.2. Pearson's correlation coefficients r values for the 15 FAO soil predictors and the five PCA soil axes. Values in italic r>0.4 or r<-0.4 (modest correlation) and in bold italic r>0.7 or r<-0.7 (strong correlation). CEC = Cation Exchange Capacity; C:N = Carbon:Nitrogen.

	PCA01	PCA02	PCA03	PCA04	PCA05
Base saturation % topsoil	-0.1013 ns	***-0.8429*** ***	0.0974 ns	0.2864 ns	0.2268 ns
CEC clay topsoil	*0.5712* ***	-0.3342 *	*-0.5284* ***	0.4161 *	-0.1366 ns
CEC soil topsoil	***0.7449*** ***	-0.1708 ns	0.2366 ns	0.0295 ns	0.0806 ns
C:N-ratio class topsoil	*0.5083* ***	0.3100 ns	*-0.4183* **	-0.2982 ns	*0.5314* ***
Easy available water	***-0.7886*** ***	*0.4747* **	0.0332 ns	-0.1926 ns	0.1242 ns
Effective soil depth	0.2428 ns	-0.3322 *	-0.1498 ns	***-0.8224*** ***	-0.2733 ns
Nitrogen % topsoil	***0.7360*** ***	0.2317 ns	0.3555 *	0.1245 ns	-0.1529 ns
Organic carbon % topsoil	*0.5523* ***	*0.5221* ***	0.3205 *	0.0227 ns	-0.2646 ns
Organic carbon pool	***0.7626*** ***	0.3883 *	0.2412 ns	-0.1427 ns	0.1172 ns
pH topsoil	*-0.4389* **	*-0.6870* ***	0.1403 ns	0.2410 ns	-0.0953 ns
Soil drainage class	***0.8323*** ***	-0.2111 ns	0.1628 ns	0.1071 ns	0.2241 ns
Soil moisture storage capacity	***-0.7108*** ***	*0.5545* ***	-0.0116 ns	-0.0222 ns	0.1731 ns
Soil production Index	-0.0489 ns	***-0.8584*** ***	0.0836 ns	-0.1758 ns	-0.2444 ns
Textural class subsoil	-0.2891 ns	-0.0161 ns	***0.8747*** ***	-0.1090 ns	0.1550 ns
Textural class topsoil	-0.0382 ns	-0.1762 ns	***0.9153*** ***	-0.1784 ns	0.1307 ns

*** p<0.05; ** p<0.01; *** p<0.001; ns = not significant.**

predictors (FAO, 2002) and elevation ranges per grid cell based on the 90m resolution SRTM altitude data (<http://srtm.csi.cgiar.org>) were added. Finally, a data-layer reflecting the El Niño Southern Oscillation (ENSO) drought impact, defined as the relative average annual difference in Normalized Difference Vegetation Index (NDVI) values between months of a severe ENSO (07/1982-06/1983) and non-ENSO years (07/1981-06/1982) was added. We retained only grid cells with data for all environmental predictors, resulting in 8577 grid cells for Borneo. Records on the coastline falling just outside grid cells were shifted to their closest grid cell. Data-layer manipulations were performed with Manifold GIS (Manifold Net Ltd).

The inclusion of 37 environmental predictors introduces potential problems of multicollinearity (Graham, 2003), which results in SDM over-fitting (Peterson *et al.*, 2007). To prevent over-fitting, the number of environmental predictors was reduced. From the DEM, elevation range, ENSO and the 19 bioclimatic predictors, we selected the least correlated variables (Maximum Pearson's r = 0.737): DEM (correlated with and proxy for elevation range, and strongly negatively correlated with the mean annual temperature, maximum temperature warmest month and quarter, minimum temperature coldest month and quarter, mean temperature wettest and driest quarter); Bio04 - Temperature seasonality; Bio07- Temperature annual range (correlated with diurnal temperature range); Bio12 - Annual precipitation (correlated with and proxy for precipitation in the wettest month and quarter, driest month and quarter, warmest and coldest quarter); Bio15 - Precipitation seasonality; and ENSO (Table 5.1). To reduce the number of soil predictors we used a principal component analysis (PCA). The PCA was performed on the 41 unique combinations of the 15 soil predictors values observed for the 8577 grid cells of Borneo, to prevent overweighting of combinations overrepresented in the dataset. We selected the first five PCA-axes as our soil property predictors PCA01-05, together describing 83% of total variance in soil data. Pearson's correlation was used to determine which of the 15 FAO soil predictors were significantly correlated with the five PCA axes (Table 5.2). This reduced environmental predictors from 37 to 11 uncorrelated variables, which were used to construct SDMs (Table 5.1).

Presence/Absence matrix from significant SDMs

To model species distributions, we selected Maxent from the available suite of applications (ver. 3.0.4; <http://www.cs.princeton.edu/~shapire/maxent/>) (Elith *et al.*, 2006; Phillips *et al.*, 2006; Graham *et al.*, 2008). Maxent was set to use all species presence records for model building, by setting the 'random test percentage' to zero (see below). The following modelling rules were used: linear features for <10 records; adding quadratic features for 10-14 records; finally including hinge features for ≥15 records (Raes & ter Steege, 2007 - Chapter 3). For all 2273 species in the dataset, an SDM was developed based on presence records and the 11 environmental predictors.

To test the significance of the SDMs, we used the bias corrected null-model method of Raes & ter Steege (2007). This method tests whether an SDM's AUC value - a threshold independent and prevalence insensitive measure of model accuracy (Fielding & Bell, 1997; McPherson *et al.*, 2004; Raes & ter Steege, 2007 - Chapter 3) - is significantly different from expectations under random chance, taking into account the uneven distribution of collection localities. Advantages of this test are that a) it allows to make use of all presence records for SDM building and testing, b) the test corrects for bias in collection localities, and c) it allows significance testing of the SDM (Raes & ter Steege, 2007 - Chapter 3). One drawback of this methodology is that habitat generalists with a wide distribution will likely not deviate from random chance expectation. It is questionable, however, to what extend these species shape floristic regions. Only species with significant SDMs were retained in subsequent floristic analyses.

Constructing an extrapolated species presence/absence matrix from the continuous Maxent SDMs required setting a threshold for each significant SDM. Although species identifications, and georeferencing of the collection localities, were carried out with the greatest possible accuracy, we assumed that 10% of the records were either wrongly georeferenced or misidentified. Therefore, for all significant SDMs represented by ≥ 10 records, the fixed '10 percentile presence' threshold was used. For those with <10 records, we used either the 'sensitivity-specificity equality' or the 'sum maximization' threshold (Liu *et al.*, 2005), dependent on which of the two corresponding omission rate values was closest to 10%. This procedure allowed us to develop the presence/absence matrix of all species with a significant SDM for the 8577 grid cells covering Borneo.

Cluster analysis and delineation of the floristic regions

We selected a hierarchical cluster analysis to group the data in floristic regions with Sørensen's index as a distance measure (Proches, 2005) in combination with the 'flexible beta linkage method' (ß = -0.25). This combination is recommended because it is space-conserving, thereby avoiding distortion, and it has the least propensity to 'chain' (McCune & Grace, 2002; Perrin *et al.*, 2006), i.e. the tendency to link entities together through a series of intermediates to form large clusters and thereby fail to find distinct cluster groups.

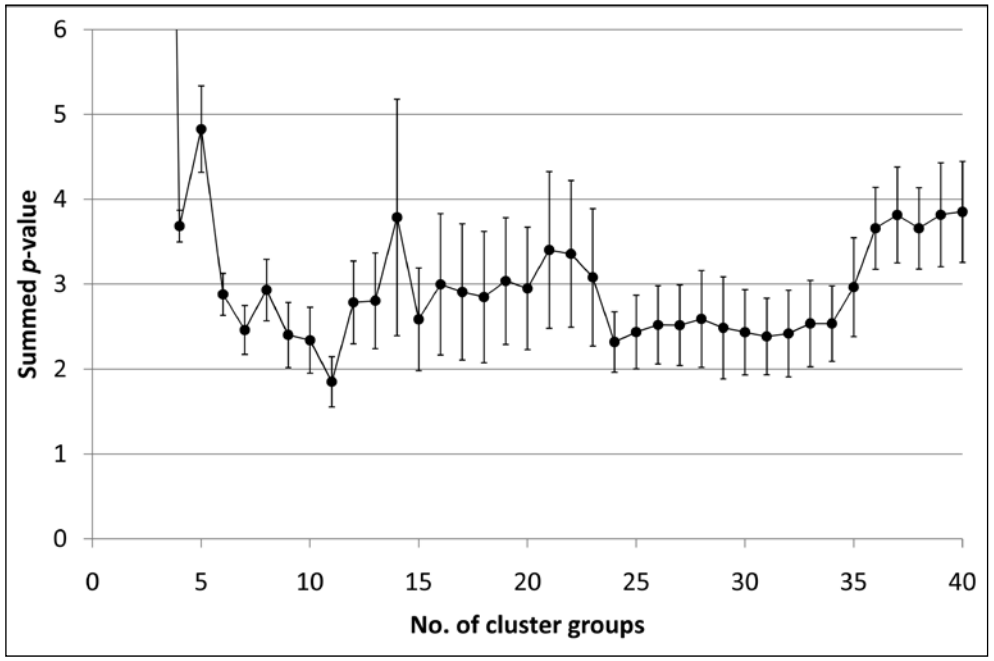


Figure 5.1. Average summed *p*-values (± SD) of the ISA for 2-40 cluster groups based on 5 x 1000 random sub-samples from the total of 8577 grid cells of Borneo.

An objective methodology to identify the ecologically most meaningful point to prune the cluster dendrogram in order to find the optimal number of final cluster groups is the indicator species analysis (ISA) (Dufrêne & Legendre, 1997) sensu McCune and Grace (2002). Although ISA is mostly applied to quantitative species data, it can also be used on presence/absence data. When applied to presence/absence data, the first part of the indicator value (IndVal) index is modified to become the ratio of the number of species presences in a sample group to the total number of species presences. The IndVal index is maximum when a species occupies all the sites of a single cluster group only (Dufrêne & Legendre, 1997). At any given level of clustering (here 2-40 cluster groups), a species is assigned to the cluster group for which its IndVal index is maximal. The significance of the assignment to a cluster group is determined with a Monte Carlo test using 1000 randomisations (Dufrêne & Legendre, 1997; Perrin *et al.*, 2006). Finally, as an objective criterion to select the optimum number of cluster groups, we used the summed *p*-value over all species for 2-40 cluster groups. The number of cluster groups with the lowest summed *p*-value, is the level for which species are most indicative of the group to which they were assigned, hence the pruning point of the cluster dendrogram.

Following this procedure, randomization of 8577 spatially continuous grid cells with species presence-absence data derived from overlapping SDMs resulted in maximum significance at very low cluster levels (data not shown). We therefore randomly sub-sampled 1000 grid cells from the 8577 cells of Borneo, and repeated the analysis five times. For each of these five sub-samples, we calculated the summed *p*-values for 2-40 cluster groups using ISA. The number of cluster groups with lowest average summed *p*-value out of five sub-samples was taken as the most informative level of clustering, hence the pruning point to delimit the floristic regions. The indicator value (IndVal) for each species was calculated based on all 8577 grid cells to identify the most informative species per cluster group. Cluster analyses and ISA were performed in PC-Ord 5.0.

Relating floristic regions to environmental factors

Classification and Regression Trees (CART) were used as a classifier to relate environmental characteristics to resulting floristic regions derived from the cluster analysis. CART is known as a robust classifier, providing results that are relatively easy to interpret in the context of SDMs (Breiman *et al.*, 1984; Bell, 1999).

We supplied the CART algorithm with the 11 environmental variables and their uncorrelated quadratic terms to account for non-linear effects, which was the case for $PCA02^2$, $PCA04^2$ and $PCA05^2$. The range of values for all variables is given in the caption of Fig. 5.3. We used 90% of the data to derive the tree, and the remaining 10% to evaluate its predictive performance, measured by the error rate (Duda *et al.*, 2001). The cross-validation cycle of using 90% to derive a tree, and 10% to determine performance, was repeated 100 times. The full decision trees, as well as all pruned versions, were evaluated in this way, and the best model resulting from this process was retained. A given floristic region (cluster group) may appear on multiple terminal nodes of the CART, and each node thus indicates alternative environmental conditions under which that floristic region occurs (Urban *et al.*, 2002).

For each terminal node of the CART tree we assessed three aspects: 1) the percentage of the total surface of Borneo identified by the corresponding CART node; 2) the correct classification rate; and 3) the percentage of the surface area within a cluster group correctly classified by the corresponding CART node. For visual inspection, the geographical extent of each floristic region was superimposed on the areas of the corresponding terminal nodes of the CART (Urban *et al.*, 2002). This was performed for each floristic region separately. The CART analysis was conducted in MATLAB, using the function 'classregtree' from the statistics toolbox.

Results

From the 2273 modelled species, 1439 (63.3%) had a distribution pattern that differed significantly from random, and their maps were used to construct the extrapolated presence/absence matrix consisting of 1439 species x 8577 grid cells (Table S5.1).

The ISA for 2-40 cluster groups indicated that an 11-cluster group stage was the optimal pruning point of the hierarchical cluster dendrogram, as at this level the lowest average summed *p*-value was found (Fig. 5.1). The pruned dendrogram and the geographical presentation of cluster groups, representing the floristic regions, are shown in Fig. 5.2. Among these 11 groups, four 'meta-cluster' groups can be recognized: a) the central mountains covering 23.8% of Borneo's surface (cluster groups 1, 2 and 3; Fig. 5.2 and Table 5.3, '% Cluster'); b) the kerangas (4) and peat swamps (5), together covering 14.1%; c) lowland forests of South-, Central- and West Kalimantan, covering 29.9% (6, 7 and 8); and d) lowland forest of East Kalimantan (11), Sabah

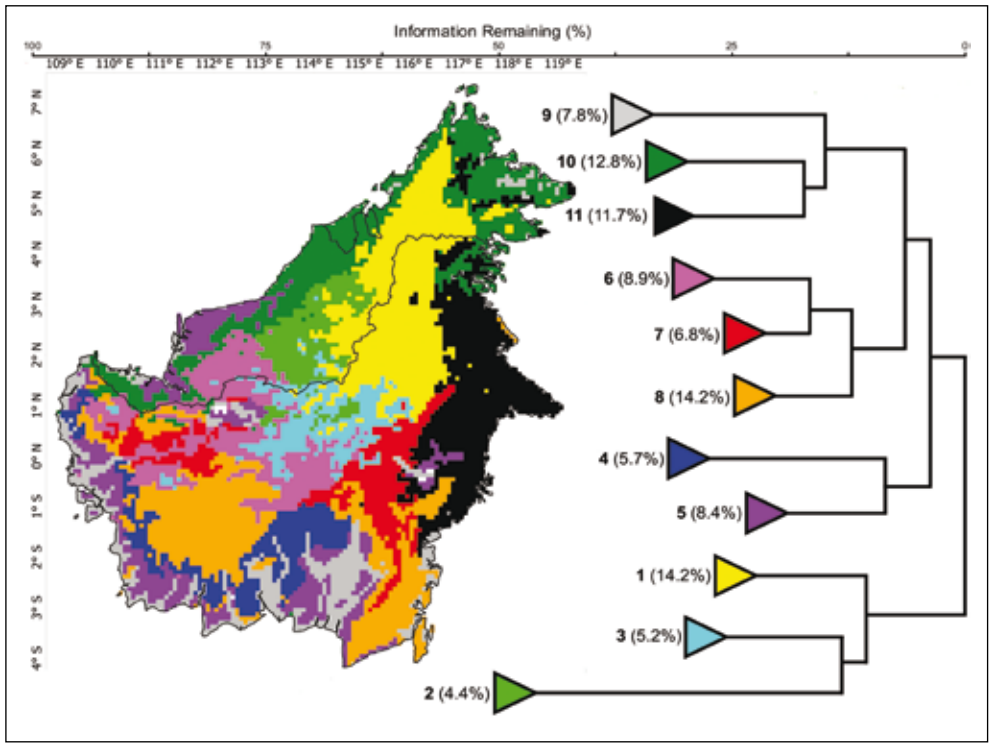


Figure 5.2. Results of the hierarchical cluster analysis representing the 11 cluster groups and their geographical distribution, indicating the different floristic regions of Borneo. Between brackets the percentage of the surface of Borneo covered by the corresponding floristic regions.

& Sarawak (10) and the fresh water swamp forests (9), together covering 32.3%. The IndVal for the 1439 significant SDMs based on the ISA of all 8577 grid cells are given in Table S5.1.

The CART, derived by selecting the tree with the smallest cross-validation error (0.076), had an overall correct classification rate of 93.1% of the cluster groups. Cluster group 3 was the less well classified, with 81.7% correct (Table 5.3). This group corresponds to the southern extrusions of the central mountain chain (Fig. 5.2; Fig. S5.1).

Generally, one terminal node -of all CART nodes classifying a given floristic region- correctly classified ≥ 50% of that region. These are given in italics in Table 5.3 and are indicated by arrows in Fig. 5.3. Floristic region 7 was an exception, with CART nodes 7.1 and 7.2 correctly classifying 33% and 40%, respectively. Most terminal nodes of the CART had a high percentage of classified cells falling within cluster groups ('% CART in Cluster'; Table 5.3). All variables were used in the CART, except PCA05[2]. The geographic representation of each terminal CART node for each floristic region is given in Fig. S5.1, along with a summation.

Discussion

Our findings demonstrate that SDMs can be used to delineate floristic regions quantitatively.

The resulting floristic map of Borneo differs most notably from previous studies by distinguishing 11 floristic regions (Fig. 5.2), compared to the six to eight regions (Fig. S5.2) (van Steenis, 1958b; Whitmore, 1984b; Wikramanayake *et al.*, 2002). The 11 floristic regions were identified using a hierarchical cluster analysis on the presence/absence matrix derived from maps of 1439 significant SDMs. This is 10% of the extrapolated 14,423 plant species expected to occur on Borneo (Roos *et al.*, 2004), including trees, lianas, shrubs and herbs.

Some of the floristic regions correspond closely to those previously recognized, but others differ in important ways, discussed below (Fig. S5.2). Most notably, the WWF defined 'Borneo lowland rain forest' ecoregion, can be divided into at least four different floristic 'sub'-regions (see also Slik *et al.* 2003). The CART correctly classified 93.1% of the grid cells based on 13 environmental variables. The most important ecological thresholds that result in the correct classification of the corresponding floristic regions are discussed per region below.

Issues concerning possible circularity of reasoning

Use of the same environmental predictors to develop the SDMs and to explain the geographical extent of the floristic regions might seem to pose a risk of circular reasoning. We argue, however, that this is not true, because Maxent identifies correlations with environmental predictors independently for each of the 2273 species. Even if for two species the same environmental predictors are used by Maxent to predict their distribution, these two species can still have other optima in their response to these predictors, hence a different spatial distribution. We therefore argue that the 1439 significant SDMs used to develop the presence/absence matrix are essentially independent (within the limited space of 11 environmental variables).

The resulting classification into 11 different floristic regions based on a cluster analysis, does not consider the environmental predictors. The reason species are predicted to co-exist in the same floristic region is not that they all occupy the same ecological niche, but rather that different niche requirements are adequately met within the same floristic region. Additionally, although the cluster tree is pruned by the optimized ISA, and cells within a floristic region are characterized by the co-occurrence of many of the same species, this does not necessarily mean that the occurrence of a given species is fully restricted to one floristic region (see Table S5.1). The CART, in its turn, identified the relevant ecological thresholds to classify the 11 floristic regions. This is an analysis at community level, derived from the cluster analysis and based on individual SDMs with their own independent response to the environmental predictors.

Results of the CART also indicate that species in the same floristic region do not all share the same ecological niche. The majority (≥ 50%) of each floristic region is classified by one CART node, generally located at the geographical centre of the floristic regions (Figs 5.2 and S5.1). At the boundaries of each floristic region, areas are found that represent CART nodes that classify smaller percentages of the floristic regions (Fig. S5.1). The smaller CART areas belonging to neighbouring floristic regions often group on the same branches of the CART tree (e.g. 1.3, 2.2, and 3.5; Fig. 5.3). These areas possibly represent areas where floristic regions are less strictly separated at the scale of our analysis, and in fact can be regarded as gradients. Where the cluster analysis puts hard boundaries, the CART indicates that edges are fuzzy.

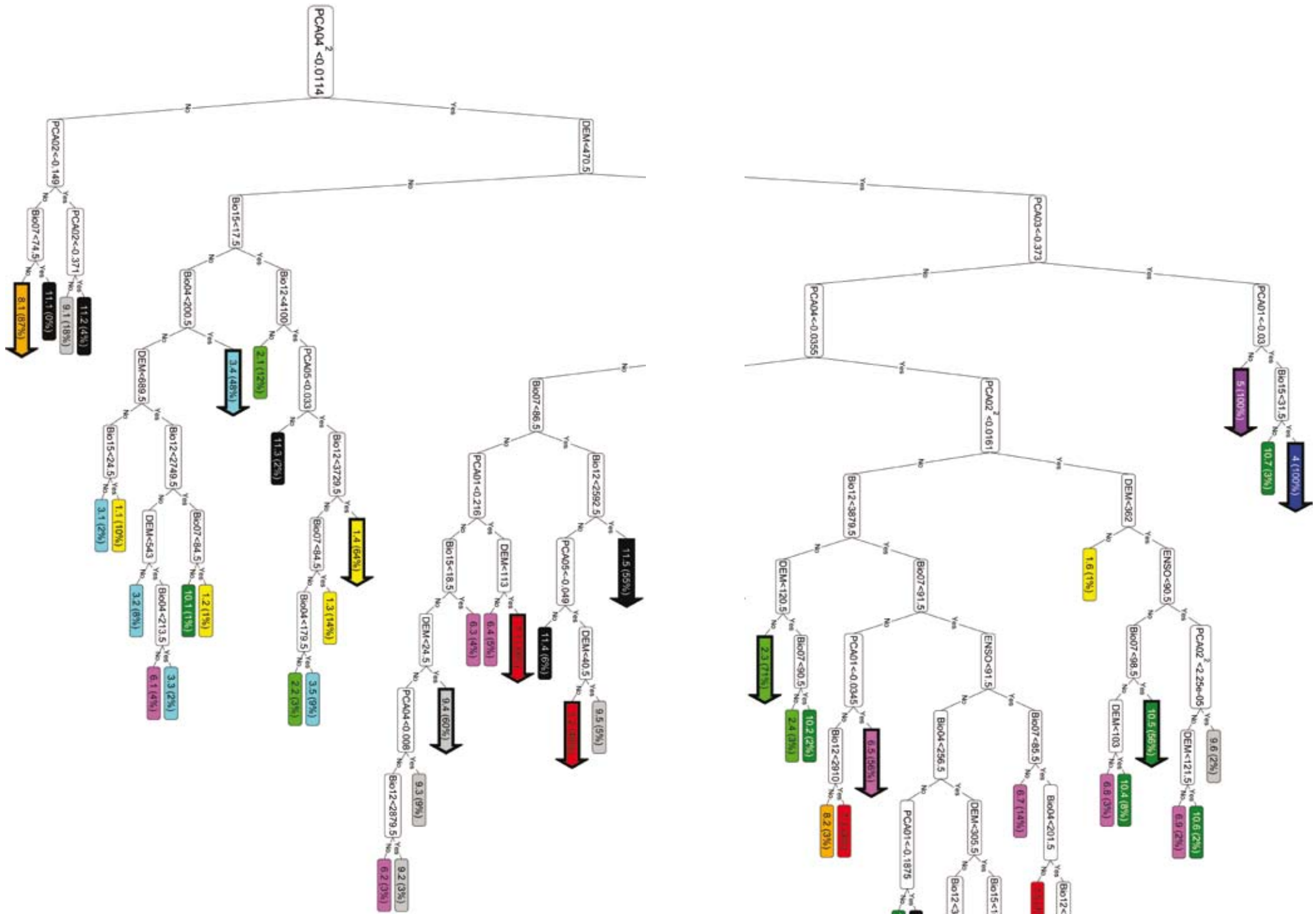


Figure 5.3. CART results based on the environmental variables and classified based on the 11 floristic regions (see Fig. 5.2). Arrows indicate the most important terminal node(s) for each floristic region. In brackets the percentage of the area covered by the corresponding floristic region correctly classified by the CART. Geographical presentations of the areas classified by each terminal node per floristic region are given in Fig. S5.1. Legend: DEM – Altitude (0 - 1972 m); Bio04 – Temperature seasonality (111 - 537 (st.dev. x 100)); Bio07 – Temperature annual range (71 - 113 (°C x 10); Bio12 – Annual precipitation (1585 - 4566 mm); Bio15 – Precipitation seasonality (9 - 67 (Coeff. of variation); PCA01 – PCA soil axis 1 (-0.505 – 0.918); PCA02 (-0.712 – 0.425); PCA022 (0 – 0.507); PCA03 (-0.67 – 0.39); PCA04 (-0.314 – 0.357); PCA042 (0 – 0.127); PCA05 (-0.246 – 0.283); PCA052 (0 – 0.080); ENSO – El Niño drought impact (81 – 108%).

Group	% Cluster	% CART	% correct classified	% CART in Cluster
01	**14.2**	**15.0**	**96.3**	**91.0**
01.1		1.4	9.8	99.2
01.2		0.1	0.7	100.0
01.3		2.4	14.0	84.2
01.4		9.5	64.4	96.1
01.5		1.4	5.9	61.5
01.6		0.3	1.4	73.9
02	**4.4**	**4.4**	**89.2**	**89.4**
02.1		0.6	11.6	91.7
02.2		0.2	3.4	81.3
02.3		3.5	70.7	89.0
02.4		0.2	3.4	100.0
03	**5.2**	**4.7**	**81.7**	**91.0**
03.1		0.2	2.5	84.6
03.2		0.5	8.1	87.8
03.3		0.1	2.5	100.0
03.4		2.6	47.7	95.9
03.5		0.6	8.5	79.2
03.6		0.3	4.5	87.0
03.7		0.5	8.1	83.7
04	**5.7**	**5.7**	**99.8**	**100.0**
05	**8.4**	**8.7**	**100.0**	**95.7**
06	**8.9**	**9.4**	**92.5**	**87.2**
06.1		0.4	3.7	82.4
06.2		0.3	3.4	86.7
06.3		0.4	4.3	86.8
06.4		0.5	4.7	76.6
06.5		5.4	55.8	92.2
06.6		0.3	1.7	56.5
06.7		1.4	14.3	90.1
06.8		0.3	2.9	88.0
06.9		0.3	1.6	42.9
07	**6.8**	**7.0**	**90.5**	**87.7**
07.1		2.5	33.0	88.1
07.2		3.0	40.0	90.3
07.3		0.2	3.1	100.0
07.4		0.2	1.7	76.9
07.5		1.0	11.3	80.5
07.6		0.1	1.4	66.7
08	**14.2**	**13.1**	**90.5**	**98.6**
08.1		12.6	87.1	98.5
08.2		0.5	3.4	100.0
09	**7.8**	**7.6**	**97.0**	**98.8**
09.1		1.4	17.6	100.0
09.2		0.2	2.7	100.0
09.3		0.7	8.7	100.0
09.4		4.7	60.5	100.0
09.5		0.5	5.1	81.0
09.6		0.2	2.4	100.0
10	**12.8**	**13.4**	**92.9**	**88.4**
10.1		0.2	1.0	61.1
10.2		0.3	2.4	89.7
10.3		2.9	20.3	90.6
10.4		1.1	8.1	93.7
10.5		7.5	56.2	95.1
10.6		0.9	1.6	21.8
10.7		0.4	3.3	97.3
11	**11.7**	**11.0**	**90.4**	**96.5**
11.1		0.0	0.0	0.0
11.2		0.5	4.7	100.0
11.3		0.3	2.4	92.3
11.4		0.8	6.7	100.0
11.5		7.1	59.7	97.9
11.6		0.7	6.2	100.0
11.7		1.5	10.8	87.2

Table 5.3. Results of the cluster analysis and the CART. '% Cluster' gives the percentage of the surface of Borneo covered by each cluster group. '% CART' gives the percentage of the surface of Borneo covered by each CART group, and for all terminal nodes separately. '% correct classified' gives the percentages of the Cluster area correctly classified by the CART group, and the percentages for all terminal nodes separately. '% CART in Cluster' gives the percentage of the surface of Borneo belonging to one CART group falling within the surface indicated by the cluster analysis, and the percentages for all terminal nodes separately.

The 11 floristic regions of Borneo

1. Montane rain forest

This largest floristic region (14.2%) covers almost the entire central mountain range, including the Crocker range and the high mountains of east Sabah, but excluding the Müller mountains in the south, which belong to region 3 (Figs 5.2 and S5.2). This region is recognized by most previous floristic classifications (Fig. S5.2). Our analysis indicates, however, that the extent is much larger than shown by Whitmore (1984) and WWF (Wikramanayake *et al.*, 2002). WWF used the 1,000 m elevation contour to delimit the lower bounds of the 'Borneo montane rain forests' ecoregion, whereas our data suggest that this floristic region extends down to 224m. One partial explanation for the lower boundary in our method is that 224m is the average altitude for a ca. 100 km^2 grid cell, but maximum altitude is much higher.

Associated areas of this region are found in valleys between the northern Crocker range, with Mt. Kinabalu, and the central mountain range and the high mountains of Sabah. Our analysis does not recognize the WWF 'Kinabalu montane alpine meadows' ecoregion as a distinct floristic region. This likely reflects the fact that in order to be modelled species had to be represented in at least five unique grid cells, whereas the 'Kinabalu montane alpine meadows' as defined by WWF occurs above 2,600 m has a surface area of only 54 km^2. Considering the spatial resolution of our analysis (ca. 100 km^2), it is not surprising that Mt. Kinabalu was not distinguished. Furthermore, the total area of the 'Kinabalu montane alpine meadows' ecoregion as reported by WWF is 4,300 km^2, yet this vastly overstates the area above the lower limit of 2600m, suggesting that more than ca. 90% of the 'Kinabalu' ecoregion should, in fact, be assigned to the 'Borneo montane rain forests' ecoregion, as we have done. We do recognize, however, that Mt. Kinabalu is a separate floristic region and discuss this further below under point '12 – Miscellaneous'.

The majority of this ecoregion (64%) is characterized by areas with an altitude above 470 m, low precipitation seasonality, intermediate to low soil C:N-ratio, and less than 3,730 mm precipitation per year (Figs 5.3 - node 1.4, and S5.1 - Group 1). The intermediate to low C:N-ratio seems to contradict with the findings that montane soils have high C:N-ratio's compared to lowland (Soethe *et al.*, 2008). The low elevation areas of Borneo harbour many peat swamps, however, which have an even higher C:N-ratio. The CART indicates that the Crocker Mountains along with Mt. Kinabalu (Fig. 5.3 - node 1.1) are characterized by different ecological conditions. The precipitation seasonality is higher than most of the region and altitude is above 690m. Similar conditions are found on Gunung Mejapa, Kong Kemoel, and the northern parts of Müller Mountains in East Kalimantan (see Fig. S5.2 A).

Many of the species with their maximum IndVal for the 'Montane rain forest' floristic region (Table S5.1) belong to plant families typical of the montane rain forest, such as Clethraceae, Cunoniaceae, Ericaceae, Fagaceae, Nepenthaceae, Orchidaceae, Podocarpaceae and Symplocaceae (Whitmore, 1984a; Wikramanayake et al., 2002).

2. Wet hill forest of Sarawak

The 'Wet hill forest of Sarawak' (Fig. 5.2 - cluster 2), located south-west from the 'Montane rain forest' region (1), was not recognized as separate floristic region by WWF, MacKinnon (1997), Whitmore (1984b) or Van Steenis (1958). This region was distinguished, however, by Slik *et al.* (2003), who classified it as the 'central Sarawak' cluster. Furthermore, the Asian Regional Centre for Biodiversity Conservation (ARCBC: <http://www.arcbc.org.ph/wetlands/>) classified this region as 'The upper basin of the Baram, Tinjar and Rajang Rivers' wetland area.

The largest part of this region (Fig. 5.3 and Table 5.3 - node 2.3, 71 %) is characterized by an altitude between 120-471 m (Fig. 5.3), shallow soils (Table 5.2 - PCA04) with an intermediate to finely structured top- and subsoil (Table 5.2 - PCA03), low base saturation and soil productivity index (Table 5.2 - PCA02[2]), and annual precipitation exceeding 3,880 mm yr^{-1}, hence the classification 'wet'. Few species are fully restricted to this region, and many occur in the other two montane floristic regions (1 & 3), and in the 'Lowland rain forest of Sabah and Sarawak' (10) region (table S5.1). Species with a maximum IndVal for this region belong to montane plant families such as Rosaceae, Ericaceae, and Fagaceae; but also include species typical of lowland rain forest families such as Dipterocarpaceae, Euphorbiaceae and Annonaceae. This region can be regarded as a transition between the true montane rain forest and lowland rain forest, characterized by very wet conditions.

3. Montane rain forest of the Müller- and Upper Kapuas Mountains (incl. Gunung Saran and Gunung Niut)

The 'Montane rain forest of the Müller- and Upper Kapuas Mountains' floristic region can be regarded as an extension plus outliers of the 'Montane rain forest' region (1). The cluster analysis indicates that this region is most closely related to the 'Montane rain forest' (Fig. 5.2). The classification as a separate region largely relates to the absence of many species present in region 1 (Table S5.1). The absent species are those with the lower maximum IndVal's for region 1, meaning that many species characteristic of the 'Montane rain forest' region also occur in region 3 (Table S5.1). Only nine species had their maximum IndVal for this floristic region, largely belonging to montane plant families.

The absence of the species present in region1 may reflect the much lower collection density on southern Borneo compared to northern parts of the mountain range. When a species is modelled with Maxent on data not fully covering a species' ecological niche due to lack of collections, Maxent will not predict the occurrence of that species under those conditions. In this case, that species may be predicted absent from southern parts of the mountain chain. The opposite is probably the case for the Schwaner Mountains, with Bukit Baka and Raja, which are excluded from floristic region 3, whilst Gunung Saran and Gunung Niut are included. The Schwaner Mountains were very extensively sampled by an expedition led by H. Nooteboom in 1982-1983 including many lowland species. Considering the size of the mountains compared to the resolution of our analysis this probably resulted in the clustering of the Schwaner Mountains with lowland region 6, whilst it more likely belongs with region 3.

Five of the seven CART nodes are found on the same branches as the nodes of region 1, mainly differing in a higher precipitation seasonality and lower temperature seasonality (Fig. 5.3 - node 3.4, 48 %). This further supports the interpretation that this region should be included in the 'montane rain forest' region (1).

4. Kerangas or heath forest

The kerangas floristic region, according our study, is found in central and west Kalimantan, often located between lowland floristic regions 6, 7 and 8, and the peat swamp region (5). The kerangas, together with the peat swamp forest, are the only two regions characterized by one CART node (Figs 5.3 and S5.1) that correctly classified the entire floristic region (Table 5.3). Both kerangas and peat swamps have small values for variable PCA03, which is related to coarse textured, organic soils (Table 5.2). Kerangas differ from peat swamps in that they have lower values for PCA01, which is related to a low cation exchange capacities (CEC) and nitrogen concentration of topsoil, a small organic carbon pool, well-drained soils with a low moisture storage capacity, and little easily available water. These results accord well with descriptions of both Mackinnon *et al.* (1996) and Whitmore (1984a) that heath forests are found on 'white sand' soils derived from siliceous parent materials which are inherently poor in bases (related to the low CEC), highly acidic, commonly coarsely textured, free-drained and often covered with a thin layer of peat or humus.

Contrary to the accordance in the description of the soils on which kerangas occur, is the extent of the range of this floristic region. Generally, all authors (Whitmore, 1984b; MacKinnon *et al.*, 1996; MacKinnon, 1997; Wikramanayake *et al.*, 2002) recognize the large kerangas area in Central Kalimantan (Fig. 5.2). Additionally, they recognize smaller areas around the 'Mahakam' lakes area of East Kalimantan and northeast hereof, and around the 'Danau Sentarum' lakes of West Kalimantan. Our analysis does not classify these areas as kerangas, but as lowland forest. Again, the scale of our analysis might play a role here. Although kerangas are generally strikingly different in flora, structure and physiognomy from other forest types, under the most favourable conditions there is considerable similarity with evergreen rain forest, with dipterocarps prominent among the larger trees, and a canopy height of 27-31m (Whitmore, 1984a; Newbery, 1991; MacKinnon *et al.*, 1996). Moreover, kerangas often form part of a mosaic with other forest formations (MacKinnon *et al.*, 1996). The presence of lowland species has likely resulted in clustering of these cells with lowland floristic regions. For the same reasons our analysis does not recognize the small patches of kerangas in north-east Sarawak, indicated by Whitmore (1984b) and studied by Newbery (1991), nor the areas fringing the coast in north-east East Kalimantan indicated by WWF (Wikramanayake *et al.*, 2002) and MacKinnon (1997).

These findings are supported by the ISA (Table S5.1), which shows that many species occurring in kerangas are also found in the lowland regions. The large area indicated by MacKinnon (1997) west of the 'Danau Sentarum' lakes area in West Kalimantan was, besides for some small patches, not recognized by other authors, and is classified in our analysis as lowland rain forest region 7 (see below). Unfortunately, this region of Borneo is largely deforested (Whitmore, 1984b; Stibig *et al.*, 2007). Whether this region should be classified as kerangas therefore remains unresolved. Our study does predict additional kerangas areas inwards along the coast line of West Kalimantan. Considering the very low collection density in the area (Raes & ter Steege, 2007 - Chapter 3), this warrants further research.

5. Peat swamp forest

The 'Peat swamp forest' floristic region 5 is found along the coast of southern Borneo, between the kerangas (4) and freshwater swamp forests (9), along large parts of the coast of Sarawak, and surrounding the 'Mahakam'- and 'Danau Sentarum' lakes. Besides these areas, peat swamp forests are also found along the coast of Brunei and northeast East Kalimantan (Whitmore, 1984b; MacKinnon *et al.*, 1996; Wikramanayake *et al.*, 2002), not recognized by our study. Peat swamp forest share many species with kerangas, as well as with lowland forests (Whitmore, 1984a). For reasons similar to those provided for why some kerangas areas were not identified as such in our study, the peat swamp areas of Brunei and East Kalimantan also were not recognized. Furthermore, these two areas are located close to the northern Borneo centre of biodiversity (Ashton, 2005; Kier *et al.*, 2005), because of which our models predict the presence of many lowland species in the swamp- and lowland rain forest mosaic, hence the clustering with the lowland rain forest floristic regions.

The CART (Fig. 5.3 -node 5) characterized the occurrence of peat swamp forest by a coarse textured top- and subsoil, a high organic carbon pool, a low pH, very poorly drained, and a low soil moisture storage capacity; all in accordance with the descriptions of other authors (Whitmore, 1984a; MacKinnon *et al.*, 1996).

Many species found to be characteristic for peat swamp forests had their maximum IndVal for this floristic region, such as *Shorea albida, Copaifera palustris, Gonystylus bancanus, Combretocarpus rotundatus*; and *Dactylocladus stenostachys* and *Campnosperma coriaceum* with a maximum IndVal for lowland rain forest region 10, but also frequently found in the peat swamps (Table S5.1) (Whitmore, 1984a; Wikramanayake *et al.*, 2002).

6, 7, 8: Lowland rain forests of southern Borneo
The delineation of the lowland rain forest floristic regions 6, 7 and 8 are the least clear-cut. The clustering of these three regions in one meta-cluster group (Fig. 5.2) is supported by the plot-based inventory study of lowland dipterocarp forest of Borneo by Slik ***et al.*** (2003) that classified the entire region as the 'southern cluster'. Other researchers classified these regions together with our regions 10 and 11 as lowland rain forest (Whitmore, 1984b; MacKinnon, 1997; Wikramanayake *et al.*, 2002) (Fig. S5.2). Cluster groups 6, 7 and 8 are mainly characterized by the absence of species which have their highest IndVal (Table S5.1) for lowland rain forest regions 10 and 11, and the wet hill forest of Sarawak (2) . These findings could support the suggested presence of the savanna corridor at the last glacial maximum (LGM ~21,000 BP) (Heaney, 1991; Gathorne-Hardy *et al.*, 2002; Bird *et al.*, 2005), when obligate rain forest species were retracted to refugia, probably present in northern and eastern Borneo (Ashton, 1992; Gathorne-Hardy *et al.*, 2002). For many species occurring in the Australian Wet Tropics (Graham *et al.*, 2006), and Europe (Svenning *et al.*, 2008), it was shown that they still have not occupied their potential range, and many are restricted to their Pleistocene refugia even today. The same may account for the lowland species of Borneo. An alternative explanation is the much lower collection density for southern Borneo compared to the north, as was explained for region 3.

Region 6
The majority of region 6 (Figs 5.3 - node 6.5, 56% and S5.1) is closely related to the 'Wet hill forest of Sarawak' (2), differing in that it receives less than 3,880 mm precipitation yr^{-1}. On the same branches of the CART tree is the most important node characterizing region 10 (10.5) found, which is different from regions 6 and 2 in that it has intermediate values for PCA02 (Table 5.2), related to base saturation, organic carbon content, pH, soil moisture storage capacity, and the soil production index. The ISA (Table S5.1) indicates that indeed many species are found in all three floristic regions. Only nine species had their maximum IndVal for region 6, and all these species had part of their distribution in other floristic regions as well.

Region 7
This region is distributed over two large geographically disjunct regions (a property shared with region 8). It is the only region, however, that is characterized by two equally imported CART nodes (Fig. 5.3 - node 7.1 and 7.2). The eastern part, node 7.2 is ecologically most closely related to the 'Lowland rain forest of East Kalimantan' (11) (Fig. 5.3 - node 11.5), different in that it receives more than 2,593 mm precipitation yr^{-1}. The western part, node 7.1, is ecologically most similar to the 'Fresh water swamp forest' (9) different in that it has lower values for PCA01 (Fig. 5.3 - node 9.4). Low PCA01 values were also the condition separating kerangas from peat swamps. This could indicate that this region has similarities with the kerangas, possibly explaining why MacKinnon (1997) characterized the area as such (Fig. S5.2 C). The threshold for PCA01<0.216 is higher than for the kerangas (PCA01<-0.03), however. This western region might be a distinct floristic unit with elements of 'Kerangas' (4) and 'Fresh water swamp forest' (9); some isolated kerangas elements on sandstone cuesta formations have been reported (G. Paoli, pers. comm.). Unfortunately, only very small patches remain forested today (Stibig *et al.*, 2007).

Region 8
Although MacKinnon (1997) characterized the region as lowland rain forest, he did recognize the 'Meratus Mountains of South Kalimantan' and the 'southern lowland planes' as separate biogeographic units (Fig. S5.2 C - 25i and 25b, respectively). The CART indicates that region 8 is characterized by partly shallow, and partly deep soils (Fig. 5.3 - $PCA04^2$>0.0114), low in base saturation, soil production index and pH (Fig. 5.3 - PCA02<-0.149). Although the floristic similarity of the two parts of region 8 is supported by similarities in our soil data (FAO, 2002), there are indications that the geological origin of both areas is different. The western part of the region 8, located in Central Kalimantan, consists of 'Schwaner Cretaceous granites & tonalites' (Hall *et al.*, 2008), and was characterized by MacKinnon (1996) as floristically relatively poor based on lithology. The eastern 'Meratus Mountains' area of region 8, together with the southern extension of the eastern part of region 7, is believed to be a cretaceous accretionary complex (Wakita & Metcalfe, 2005), characterized by Hall *et al.* (2008) as a mixture of the 'Meratus Cretaceous volcanic arc & ophiolite' and 'Ophiolite basic/ultrabasic rocks'. Whether these recent discoveries are correctly reflected in our soil property variables (Table 5.2) is not known.

Like for region 6 and 7, the ISA indicated that few species had their maximum IndVal for region 8 (Table S5.1). That the Meratus Mountains are not recognized as a distinct floristic region might be for reasons similar to those offered for Mt. Kinabalu, namely study resolution and low number of species with a minimum of five presence records (see region 1, above). According to MacKinnon (1996), the Meratus Mountains are floristically distinct and especially rich in orchids.

9. Fresh water swamp forest
The 'Fresh water swamp forest' is, not surprisingly, found in the same cluster as the lowland rain forest regions 10 and 11 (Fig. 5.2), since this floristic region has many species in common with the surrounding lowland rain forest (Whitmore, 1984a; Wikramanayake *et al.*, 2002), as can be concluded for the ISA (Table S5.1). Many species belonging to the families of the Cyperaceae and Rhizophoraceae have their maximum IndVal for this floristic region. The occurrence of the rain- and river fed swamp forest up to the coast has resulted in the lumping of the 'Fresh water swamp forest with the mangrove forest. The areas identified as 'Fresh water swamp forest' are largely in accordance with the findings of other authors (Fig. S5.2). The only areas not recognized by our study are located in northern East Kalimantan.

The CART (Fig. 5.3) indicates that the most important node 9.4, classifying the fresh water swamps, is found on the same branch system as the node 11.5 classifying the 'Lowland rain forest of East Kalimantan' which supports the floristic similarities between the regions. Node 9.4 is related to high CEC's, high nitrogen percentage, and poorly drained soils with high soil moisture capacity (Table 5.2 - PCA01>0.216), all in accordance with the descriptions of Whitmore (1984a).

10. Lowland rain forest of Sabah and Sarawak
Contrary to previous floristic analyses of Borneo (Whitmore, 1984b; MacKinnon, 1997; Wikramanayake *et al.*, 2002) that classified the Borneo lowlands as one region, our analysis indicates that the 'Lowland rain forest of Sabah and Sarawak' are distinct from lowland regions 6, 7 and 8, and 11 (Fig. 5.2). These results are supported by Ashton (1992), who labelled the western, and north-western parts of Sarawak as areas of exceptional richness and endemicity, the so-called Riau-pocket (Corner, 1960; Ashton, 1992). The north-eastern part of region 10 was identified by both Slik *et al.* (2003) and Ashton (1992) as the separate 'East coastal Sabah sub-province', however. Our study shows that the north-eastern part of Borneo is a

mosaic of lowland rain forest regions 10 and 11, fresh water swamp forest, and montane rain forests, and can be described as an area where the 'Lowland rain forest of Sarawak and Sabah' and the 'Lowland rain forest of East Kalimantan' mix.

The CART characterized the area by deep soils (Fig. 5.3 - PCA04<-0.0355), which is supported by the presence of the Crocker Fan consisting of Tertiary basin fills in north-western Sarawak, and Tertiary and Quaternary sediments in west Sarawak and in Sabah (Hall *et al.*, 2008). The CART further indicates that this region is least struck by ENSO droughts and is found below 362 m altitude (Fig. 5.3 - node 10.5). The ISA (Table S5.1) found most species with a maximum IndVal for this floristic region, which further support the findings of Ashton (1992) that the region is characterized by exceptional richness and endemicity.

11. Lowland rain forest of East Kalimantan

Although the 'Lowland rain forest of East Kalimantan' is most closely related to the 'Lowland rain forest of Sabah and Sarawak', they are two distinct floristic regions and not one, as suggested by Slik *et al.* (2003). The ISA (Table S5.1) indicates that there are many species characteristic for this region, which is further supported by the presence of Ulin (*Eusideroxylon zwageri*) in the lowland forests of East-, South- and Central Kalimantan but not in Brunei, Sabah and Sarawak (van Steenis, 1958b). That the lowland forests of the Kalimantan provinces are different from those of Sabah and Sarawak was also recognized by Ashton (1992), who classified the entire region as 'The generalized inland flora on udult ultisol soils'.
The CART indicated that the majority of the region has shallower soils than region 10 (Fig. 5.3 - PCA04>-0.0355), has a small annual temperature range (Bio07<8.7 °C) and receives less than 2,593 mm precipitation year^{-1}. The many species with their highest IndVal (Table S5.1) for this region support the separate floristic status. Although the CART paths to this region did not include the ENSO drought predictor, it is this region of Borneo which is most severely struck by ENSO droughts, an important reason why large parts of this region are deforested (Langner *et al.*, 2007).

12. Miscellaneous – Not distinguished

There are several additional floristic regions present on Borneo not classified as such by our analysis. These include: the Kinabalu montane meadows, discussed under 1; the Meratus Mountains, discussed under 8; the mangroves (Whitmore, 1984a, b), classified in our analysis with freshwater swamps (9); and the limestone and ultramafic floristic regions not included because of a lack of data. The exact extent and location of these latter two regions is not clear, a reason why WWF did not recognize these in their ecoregions (Wikramanayake *et al.*, 2002). We do recognize the existence of these floristic regions, but mainly due to resolution of our analysis we were not able to distinguish them.

Conclusions

For the first time, the floristic division of Borneo is quantitatively analysed based on 1439 significant SDMs and a cluster analysis that has resulted in the recognition of 11 floristic regions. Some doubts remain, however, especially about the southern Borneo lowland rain forest regions. These regions are largely characterized by the absence of indicator species, which could support the presence of a savanna corridor present during the LGM, forcing obligate rain-forest species to the northern refugia from where they still not have recovered their potential distribution range, a phenomenon known as the postglacial migration lag (Svenning *et al.*, 2008). Another issue that could have resulted in a different floristic delineation of southern Borneo is related to the fact that the most common species are excluded from our analysis based on statistical grounds. Common species, which show no niche preference, cannot be distinguished from a random null-model.
The presence of common species throughout all lowland areas results in more similar Sørensen's index values, potentially resulting in the aggregation of cluster groups 6, 7 and 8. On the other hand, the deficiency of collections from eastern, and especially southern Borneo (Raes & ter Steege, 2007 - Chapter 3), potentially has resulted in less recognized endemic species for these regions, since the limit for modelling was set at five records. The chance that an endemic from these regions is represented by five records is therefore much lower. Recognition of potential endemic species could have resulted in more clearly distinct cluster groups with their own indicator species. The same arguments could account for the relative absence of indicator species for the 'Montane rain forest of the Müller- and Upper Kapuas Mountains' (3) compared to the 'Montane rain forest of the central mountain chain' (1). To clarify these issues additional collections are required.

Nonetheless, our study quantitatively supports most of the previously recognized floristic regions, while at the same time indicates that the lowland rain forest of Borneo, often regarded as one floristic region (Fig. S5.2), should be divided in at least four (and potentially six) different regions: The lowland rainforests of 'Sabah and Sarawak'(10), 'East Kalimantan' (11), and 'southern Borneo' (6, 7, 8), and the 'Wet hill forest of Sarawak' (2).

Acknowledgements - We like to thank Jaap Vermeulen for sharing his orchid database, Ben Kieft for digitizing all maps, Gary Paoli for many constructive comments, the Alberta Mennega Stichting for funding the colour images, and the NWO-Groot 'Building databases for life'-project (Grant# 175.010.2003.010) for funding the digitization of all Borneo collections used for this research.

CHAPTER 6

Borneo's remaining forests – Where to from here?

Niels Raes, Pieter Baas, E. Emiel van Loon, Marco C. Roos, J.W. Ferry Slik and Hans ter Steege

In preparation

In this chapter we summarize the results of the previous chapters, show how much the botanical diversity and floristic regions of Borneo have been impacted, and discuss the conservation/policy implications of these results. Finally, we make suggestions for future research to answer questions raised as result of our studies, and to further improve our understanding of what shapes macroecological biodiversity patterns. Although it is widely recognized that Borneo is one of the world's most important biodiversity hotspots (Myers *et al.*, 2000), the spatial patterns of botanical richness, endemicity, 'centres of endemicity', and Borneo's floristic regions, have until now largely been based on informal expert opinion. Recent digitization of the botanical collections of Borneo, housed at the National Herbarium of the Netherlands, has provided a database that allowed a quantitatively spatial analysis of the components of biodiversity of Borneo. We have shown that botanical richness and endemicity are not evenly distributed over Borneo, and also that clear floristic regions, with meaningful ecological correlates can be identified. Much of Borneo has been impacted by man, however (Curran *et al.*, 2004; Dennis & Colfer, 2006; Langner *et al.*, 2007; Stibig *et al.*, 2007), but to what extent this has affected the species rich areas, the centres of endemicity and the various floristic regions remains unknown.

The 'road map' to patterns of botanical diversity and the floristic regions of Borneo

We selected species belonging to families treated in Flora Malesiana (Anon., 1959-2007), together with those of the revised genera of the Annonaceae, Euphorbiaceae, and Orchidaceae. To georeference the collections we used online- and printed gazetteers. Mapping of the georeferenced collections revealed an uneven, or biased, distribution of collection localities on Borneo (Fig. 3.3). To reduce bias to a minimum we applied a technique known as georegistration in order to georeference as many collections as possible from the severely under-collected provinces of Kalimantan (Raes *et al.*, 2009 - Chapter 2). Our efforts resulted in a database comprising 66,262 georeferenced records belonging to 102 plant families representing 2,273 species.

To develop Borneo-wide biodiversity patterns at high spatial resolution we used a technique known as species distribution modelling. Species distribution models (SDMs) predict the potential distribution of a species by describing relationships between a species' presence/absence-, or presence-only data, and a set of environmental predictors across an area of interest. Depending on the availability of meaningful environmental predictors, in combination with sufficiently accurate collection localities, SDMs can predict the presence and absence of species across the entire area of investigation at the spatial resolution of the environmental predictors (Guisan & Zimmermann, 2000; Araújo & Guisan, 2006; Peterson, 2006). From the available suite of modelling applications (Elith *et al.*, 2006; Pearson *et al.*, 2006) we selected Maxent (ver. 3.0.4) (Phillips *et al.*, 2006) for our data, because Maxent was a) specifically developed to model species distributions with presence-only data (typical of herbarium data), b) has shown to outperform most other modelling applications (Elith *et al.*, 2006; Pearson *et al.*, 2007; Wisz *et al.*, 2008), and c) is least affected by georeferencing errors (Graham *et al.*, 2008).

The application of predictive models, like SDMs, require testing of their predictive accuracy. We showed, however, that currently used SDM accuracy measures based on presence-only data, and pseudo-absences instead of true absences (which are often not available), cannot reliably be applied (Raes & ter Steege, 2007 - Chapter 3). Therefore, we introduced a newly developed null-model methodology that tests whether an SDM's AUC value - a threshold independent and prevalence insensitive measure of model accuracy (Fielding & Bell, 1997; McPherson *et al.*, 2004; Raes & ter Steege, 2007 - Chapter 3) - is significantly different from random chance expectation, taking into account the uneven distribution of collection localities.

We developed SDMs for the 2273 species based on their presence records, and 11 meaningful and independent environmental predictors at 5 arc-minute (ca. 100 km^2) spatial resolution (Raes *et al.*, submitted - Chapter 4). All models were tested against a bias corrected null-model, resulting in 1439 significant SDMs (63.3%), covering 8577 grid cells. We converted the continuous Maxent predictions to discrete presence/absence maps by applying a 10-percentile threshold. Significant SDMs were superimposed to generate the botanical richness pattern of Borneo (Fig. 6.1A). As measure of endemicity we used the weighted endemism index (Crisp *et al.*, 2001; Kier & Barthlott, 2001; Küper *et al.*, 2006; Slatyer *et al.*, 2007). We developed the endemicity pattern by summing the weights of all 1439 significant SDMs for all grid cells (Fig. 6.1C). The 'centres of endemicity' were identified by mapping the relative residuals of the species richness – weighted endemism relationship (Fig. 6.1E).

The 50 percent highest diversity grid cells cover 15.8% of Borneo (Table 6.1, quartile 3 & 4); for weighted endemism this is less than 5%. Only 2.1% of Borneo's surface has more than 150% endemics than can be expected based on their diversity values (Table 6.1, '50-100' & '>100'). These areas are the 'centres of endemicity'. The areas of high diversity and endemicity are characterized by a relatively small range in annual temperature, but with seasonality in temperatures within that range. Furthermore, these areas are least affected by the El Niño Southern Oscillation drought events. The 'centres of endemicity' are found in areas that are ecologically distinct in altitude, edaphic conditions, annual precipitation, or a combination of these factors.

To identify the floristic regions of Borneo we constructed a presence/absence matrix based on 1439 significant SDMs for the 8577 grid cells of Borneo (Raes *et al.*, submitted - Chapter 5). This matrix was then analysed using a hierarchical cluster analysis, and the resulting cluster dendrogram was pruned using indicator species analysis (ISA) to partition the 11 floristic regions (Fig. 6.1G; Table 6.1, Floristic Region). The relationship between the 11 floristic regions and environmental conditions was explored using a classification and regression tree (CART) analysis. CART identified meaningful ecological thresholds defining each floristic region, largely in accordance with the known ecology of the represented 'forest types' (Whitmore, 1984a; Wikramanayake *et al.*, 2002). This method allowed the quantitative confirmation of the floristic distinctiveness and extent of montane rain forest (Floristic Region 1 & 3), kerangas (4), peat swamps (5), and fresh water swamp forest (9). The lowland rain forest, previously recognized as one floristic region (Whitmore, 1984b; MacKinnon, 1997; Wikramanayake *et al.*, 2002) was divided in at least four (and possibly six) distinct floristic regions, viz. the lowlands of (i) Sabah and Sarawak (10), (ii) East Kalimantan (11), (iii) southern Borneo (6, 7 & 8), and (iv) the Wet hill forest of Sarawak (2). Due to the 100 km^2 resolution of our analysis, we could not distinguish, but do recognize, the 'Kinabalu highlands', mangroves, and forests on limestone and ultramafic rock.

Borneo's remaining forests

The use of SDMs to generate quantitative patterns of botanical richness, weighted endemism, 'centres of endemicity', and floristic regions of Borneo, resulted in potential forested extents of these areas (Fig. 6.1, 'Total'), with 100% of Borneo covered by forest. Much of Borneo has been impacted by man, however, and few areas have been put aside in conservation areas. To analyse the extent of deforestation and conservation in areas of high diversity, high endemicity and the different floristic regions, we overlaid our maps with those of forest change and conservation areas. As a proxy for Borneo's remaining forests we used the Global Land Cover 2000 dataset (GLC2000, 2003) for South East Asia

Table 6.1. The percentages of Borneo's surface covered by the four quartiles of 'Species Diversity' and 'Species Weighted Endemism', the four 'Relative Residual Weighted Endemism' classes, and the 11 'Floristic Regions' for its entire surface - 'Total', divided in 'Forested' and 'Non-Forested' extents, and for 'Protected-' and 'Non-Protected Areas' divided in 'Forested' and 'Non-Forested' extents. Between brackets percentages deforestation.

		Total	Forested	Non-Forested	Protected Area		Non-Protected Area	
					Forested	Non-Forested	Forested	Non-Forested
Species	**1**	62.4	40.3	22.1 (35.4)	5.7	0.9 (13.6)	34.6	21.2 (38.0)
Diversity	**2**	21.7	14.0	7.7 (35.3)	1.1	0.4 (28.0)	12.9	7.2 (35.9)
	3	10.2	6.8	3.4 (33.7)	0.7	0.3 (27.8)	6.1	3.2 (34.3)
	4	5.6	2.4	3.2 (57.4)	0.4	0.2 (33.0)	2.1	3.1 (59.9)
Species	**1**	81.3	52.3	28.9 (35.6)	6.4	1.3 (16.9)	45.9	27.6 (37.6)
Weighted	**2**	14.5	9.2	5.3 (36.6)	1.0	0.3 (24.4)	8.2	5.0 (37.9)
Endemism	**3**	3.3	1.4	1.9 (56.8)	0.2	0.1 (32.2)	1.2	1.8 (59.6)
	4	0.9	0.6	0.3 (35.7)	0.2	0.0 (14.3)	0.4	0.3 (43.7)
Relative	**<0**	59.8	36.5	23.3 (39.0)	2.5	1.1 (30.5)	34.0	22.2 (39.5)
Residual	**0-50**	38.0	25.2	12.8 (33.6)	4.5	0.6 (12.3)	20.7	12.2 (37.0)
Endemism	**50-100**	1.9	1.6	0.3 (16.6)	0.7	0.0 (4.4)	0.9	0.3 (24.8)
(in %)	**>100**	0.2	0.2	0.0 (5.9)	0.1	0.0 (3.5)	0.1	0.0 (7.6)
Floristic	**1**	14.3	13.1	1.2 (8.5)	2.4	0.1 (5.2)	10.7	1.1 (9.2)
Region	**2**	4.5	4.2	0.3 (7.2)	0.0	0.0 (0.6)	4.1	0.3 (7.3)
	3	5.3	5.2	0.1 (1.5)	1.5	0.0 (1.7)	3.6	0.1 (1.4)
	4	5.8	3.1	2.6 (45.7)	0.6	0.1 (15.3)	2.6	2.5 (49.6)
	5	8.3	4.5	3.8 (46.3)	0.6	0.2 (26.6)	3.8	3.6 (48.6)
	6	9.0	6.1	2.9 (32.2)	0.9	0.2 (14.7)	5.2	2.7 (34.5)
	7	6.9	3.4	3.5 (50.5)	0.0	0.0 (70.7)	3.4	3.4 (50.4)
	8	14.3	7.9	6.4 (44.7)	0.8	0.2 (24.1)	7.1	6.1 (46.2)
	9	7.4	3.8	3.7 (49.6)	0.2	0.1 (32.7)	3.5	3.6 (50.3)
	10	12.6	6.9	5.6 (44.9)	0.7	0.3 (28.1)	6.3	5.4 (46.2)
	11	11.7	5.4	6.3 (53.5)	0.1	0.5 (84.4)	5.4	5.8 (52.0)
Total		**100.0**	**63.6**	**36.4**	**7.8**	**1.8 (18.4)**	**55.7**	**34.7 (38.4)**

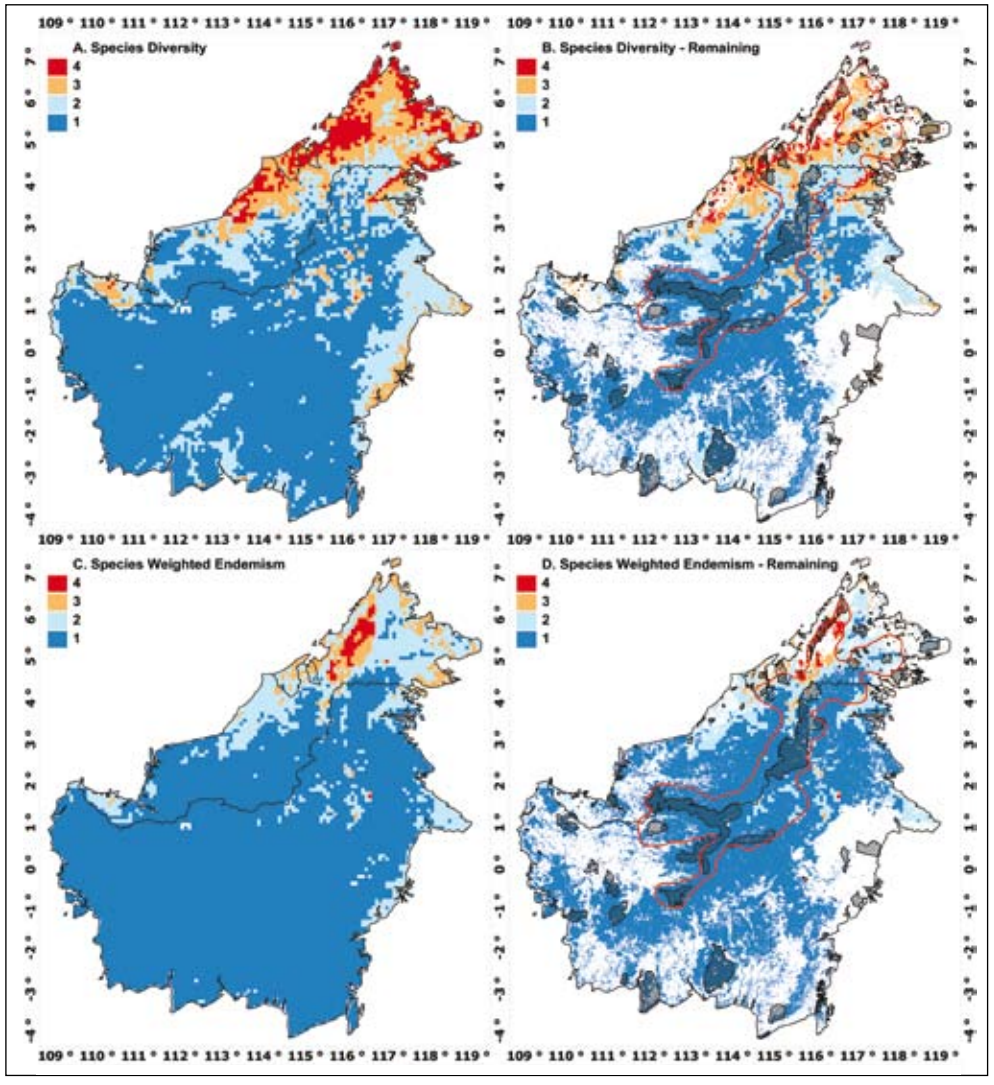


Figure 6.1. A. The four quartiles of species diversity (1=lowest diversity; 4= highest diversity); B. The four quartiles of species diversity still forested; C. The four quartiles of species weighted endemism; D. The four quartiles of species weighted endemism still forested. Red line – the proposed trans-boundary WWF 'Heart of Borneo' protected area.

based on SPOT-VEGETATION satellite data for the years 1998-2000 (Stibig *et al.*, 2007). The average annual deforestation rate on Borneo of 1.7% (Langner *et al.*, 2007) suggests that our estimates of the forested extent of Borneo are probably conservative. We re-sampled our 5 arc-minute maps to the '1 km at equator' (0.00893 decimal degree) resolution of the South East Asian land-cover map, and kept only those grid cells with values for both maps. We used land-cover forest classes 1-5 as our proxy for forested extent, and assessed the

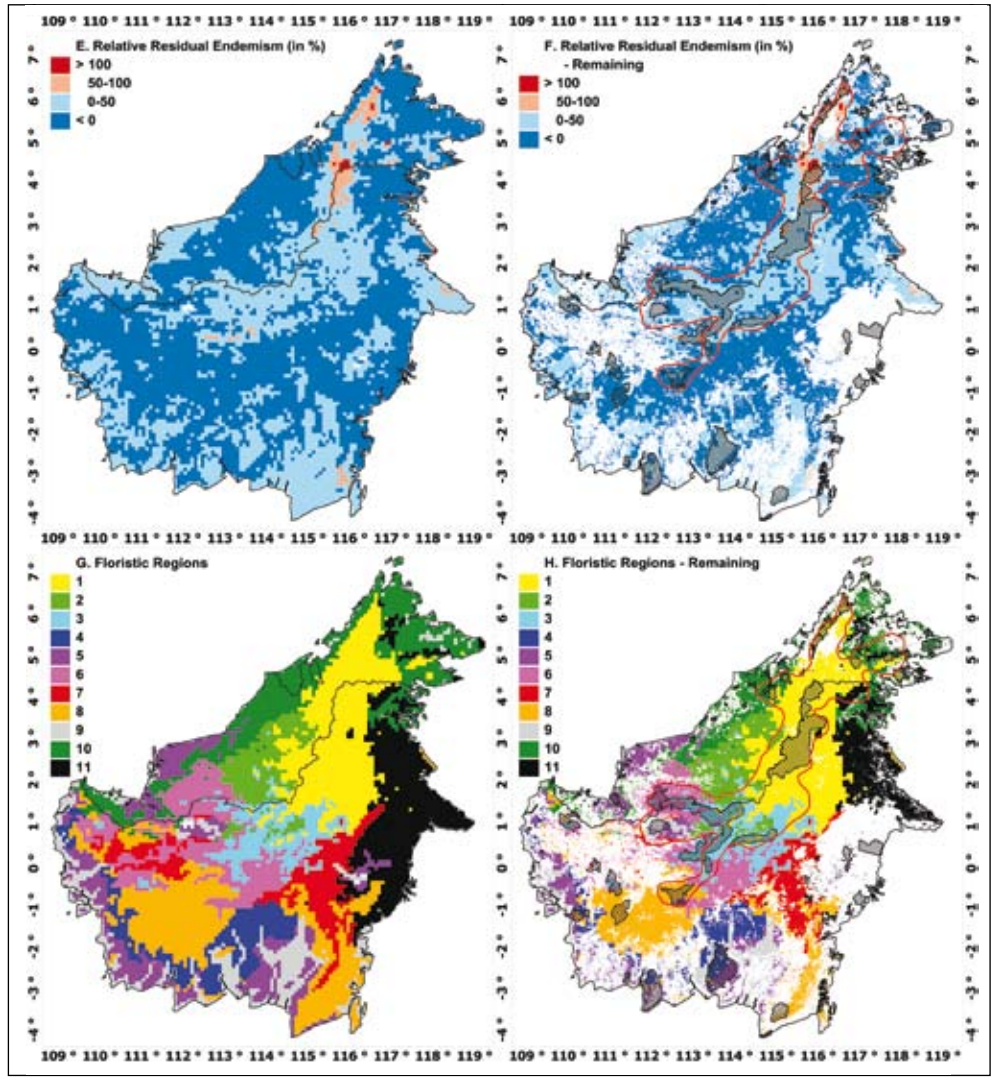


Figure 6.1. (continued) E. The four classes of relative residual weighted endemism ('<0' =less than expected, '0-50' = up to 50% more endemism than expected, '50-100' = 50-100% more endemism than expected, '>100' = more than 100% more endemism than expected); F. The four classes of relative residual weighted endemism still forested; G. The 11 floristic regions of Borneo; H. The 11 floristic regions of Borneo still forested. Hatched areas indicate the IUCN recognized protected areas (WDPA, 2007). Red line – the proposed trans-boundary WWF 'Heart of Borneo' protected area.

percentage of each area covered by these five land-cover forest classes (Fig. 6.1B, D, F, H; Table 6.1, 'Forested/Non-Forested'). The analysis reveals that 36 % of Borneo's total surface, and 57% of its most diverse areas (Table 6.1, Species Diversity - 4th quartile) are already very heavily impacted. Especially the most diverse lowlands of Sabah and Sarawak (Floristic Region 10), and those of East Kalimantan (11) have been severely hit by

deforestation caused by logging, forest fires and land-use change (Table 6.1, Fig. 6.1) (Sodhi *et al.*, 2004; Dennis & Colfer, 2006; Langner *et al.*, 2007; Stibig *et al.*, 2007).

Deforestation has taken its toll even in the IUCN recognized protected areas (Curran et al., 2004), as can be concluded from the overlay of the World Database of Protected Areas (WDPA, 2007) on the remaining forested areas (Fig. 6.1B, D, F, and H; hatched areas). Only 0.6% of Borneo's surface belongs to the fourth quartile of species diversity, while at the same time having an IUCN protected status. By the year 2000, 33% of this area was already deforested, however. For the non-protected areas belonging to this category (5.2% of Borneo's surface), 60% was already lost by 2000. For areas in the highest weighted endemism categories (Table 6.1, Species Weighted Endemism, quartile 3 & 4), values in the same order of magnitude were found. Most catastrophic is the loss of 84% of East Kalimantan's lowland rain forests in protected areas (Table 6.1, Floristic Region 11). The 'centres of endemicity' are least affected by deforestation (Table 6.1, Relative Residual Endemism), as these are mainly found in the higher altitude, less impacted, floristic regions ('Montane rain forest' – Floristic Regions 1 & 3; Table 6.1).

Implications for conservation

The latest effort to conserve large part of Borneo's biodiversity is the 'Heart of Borneo' initiative (WWF-Germany, 2005; Stone, 2007). This area of more than 20 million ha straddles Borneo's trans-boundary highlands of Indonesia and Malaysia, and reaches out through the foothills into adjacent lowlands and to parts of Brunei (Fig. 6.1 B, D, F, H; red boundary). Although this initiative is a milestone for conservation on Borneo, many floristic regions will not be protected within its boundaries (Fig. 6.1H). To safeguard Borneo's genetic diversity, especially the last remaining high diversity lowland rain forest regions of 'East Kalimantan' (11) and 'Sabah and Sarawak' (10) should be awarded protected status. For East Kalimantan, we suggest the Sungai Wain 'Protection' Forest close to Balikpapan, part of the Sangkulirang Peninsula, and the area between the Sembakung-Sesajap delta and the montane rain forest in northern East Kalimantan. For Sabah and Sarawak the areas west of the Crocker Mountains range, the valley between the Crocker Mountains and the central mountains range, and the last remaining lowlands of south-western Sarawak are suggested. Furthermore, the floristic regions that cover smaller percentages of Borneo's surface, such as 'Kerangas' (4), 'Peat swamp forests' (5), and 'Fresh water swamps' (9) should receive more conservation attention and a larger percentage of their extent should be protected. Finally, we suggest that parts of all southern Borneo lowland rain forest regions (6, 7 and 8) are conserved.

Future research prospects

The quantitative analysis of Borneo's botanical diversity, endemicity, and floristic regions has also raised several challenges that deserve to be addressed by future research.

1. The first is the introduction of error-surfaces, or the extent to which gradients in environmental predictors are covered by sample localities. These issues have started to be addressed (Kadmon *et al.*, 2004; Guralnick *et al.*, 2007; Loiselle *et al.*, 2008), but deserve additional research and incorporation in SDMs.
2. For our analyses we used single species distribution models that exclude many species on grounds of the required minimum number of presences. A recent review of community level modelling (Ferrier & Guisan, 2006) suggests that these methods can offer an approach for constructing distribution models for rare species with low occurrence data. It should be explored how results of these methods differ from our results.
3. The application of large scale analysis of combined phylogenetics and bioclimatic modelling, known as phyloclimatic modelling (Yesson & Culham, 2006b; Yesson & Culham, 2006a) should be explored to improve LGM refugial reconstructions (Waltari *et al.*, 2007), and to assess when, and under which environmental conditions speciation most likely has taken place.
4. Modelling of species traits of dispersal in relation to the post-glacial dispersal lag hypothesis (Svenning *et al.*, 2008) may lead to better founded reconstructions of LGM refugia.
5. Projection of 'local' SDMs on an SDM derived from the global set of occurrence records allows testing whether local populations occupy their entire potential niche, and whether their predictive accuracy improves, or deteriorates. Geographically separated populations can occupy a different spectrum of their niche due to genetic drift or competition with other species. Vice versa, projection of SDMs to a global environmental predictor dataset can reveal biogeographical boundaries and regions prone to invasion.
6. The combination of diversity measures from plot studies (Slik *et al.*, 2003) with biodiversity patterns derived from SDM to potentially introduce abundance measures to SDMs.
7. The projection of SDMs under different Global Climate Change scenario's (Millennium-Ecosystem-Assessment, 2005) to identify those areas which potentially can preserve most of the world's genetic diversity in a changing world.

REFERENCES

Aiba, S.-I., Kitayama, K. & Repin, R. (2002) Species composition and species-area relationships of trees in nine permanent plots in altitudinal sequences on different geological substrates of Mount Kinabalu. *Sabah Parks Nature Journal*, **5**, 7-69

Anderson, R.P., Gomez-Laverde, M. & Peterson, A.T. (2002) Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography*, **11**, 131-141

Anderson, R.P., Lew, D. & Peterson, A.T. (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, **162**, 211-232

Anon. (1959-2007) *Flora Malesiana, Series II, Pteridophyta*. Nationaal Herbarium Nederland, University Leiden branch.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677-1688

Araújo, M.B., Nogues-Bravo, D., Diniz-Filho, J.A.F., Haywood, A.M., Valdes, P.J. & Rahbek, C. (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography*, **31**, 8-15

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693-695

Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species-climate impact models under climate change. *Global Change Biology*, **11**, 1504-1513

Ashton, P.S. (1983) Dipterocarpaceae. *Flora Malesiana I* (ed. by C.G.G.J. van Steenis), pp. 237-552. Martinus Nijhoff, The Hague.

Ashton, P.S. (1992) Plant conservation in the Malaysian region. *Proc. Intern. Conf. Tropical Biodiversity "In Harmony with Nature"* (ed. by S.K. Yap and S.W. Lee), pp. 86-93. Malayan Nature Society, Kuala Lumpur.

Ashton, P.S. (2005) Lambir's Forest: The World's Most Diverse Known Tree Assemblage? *Pollination Ecology and the Rain Forest: Sarawak Studies* (ed. by D.W. Roubik, S. Sakai and A.A. Hamid Karim), pp. 191-216. Springer, New York.

Ashton, P.S. & Hall, P. (1992) Comparisons of Structure Among Mixed Dipterocarp Forests of North-Western Borneo. *The Journal of Ecology*, **80**, 459-481

Barthlott, W., Mutke, J., Rafiqpoor, D., Kier, G. & Kreft, H. (2005) Global Centers of Vascular Plant Diversity. *Nova Acta Leopoldina*, **342**, 61-83

Beaman, J.H. (2005) Mount Kinabalu: hotspot of plant diversity in Borneo. *Biologiske Skrifter*, **55**, 103-127

Bell, J.F. (1999) Tree-based methods. *Machine learning methods for ecological applications* (ed. by A.H. Fielding), pp. 89-106. Kluwer Academic, Boston, MA.

Bird, M.I., Taylor, D. & Hunt, C. (2005) Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? *Quaternary Science Reviews*, **24**, 2228-2242

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045-1055

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and Regression Trees*. Wadsworth, Pacific Grove, CA.

Buys, D.W., Witkamp, H., Endert, F.H., Siebers, H.C. & Bosch, D.F.K. (1927) *Midden-Oost-Borneo Expeditie 1925*. G. Kolff & Co., Weltevreden.

Cannon, C.H., Curran, L.M., Marshall, A.J. & Leighton, M. (2007) Long-term reproductive behaviour of woody plants across seven Bornean forest types in the Gunung Palung National Park (Indonesia): suprannual synchrony, temporal productivity and fruiting diversity. *Ecology Letters*, **10**, 956-969

Corner, E.J.H. (1960) The Malayan Flora. In: *Proceedings of the Centenary and*

Bicentenary Congress of Biology (ed. R.D. Purchon), pp. 21-24, Singapore

Costa, G.C., Nogueira, C., Machado, R.B. & Colli, G.R. (2007) Squamate richness in the Brazilian Cerrado and its environmental-climatic associations. *Diversity and Distributions*, **13**, 714-724

Crisp, M.D., Laffan, S., Linder, H.P. & Monro, A. (2001) Endemism in the Australian flora. *Journal of Biogeography*, **28**, 183-198

Curran, L.M., Trigg, S.N., McDonald, A.K., Astiani, D., Hardiono, Y.M., Siregar, P., Caniago, I. & Kasischke, E. (2004) Lowland forest loss in protected areas of Indonesian Borneo. *Science*, **303**, 1000-1003

Dalton, B. (1978) *Indonesia Handbook, Second Edition*. Moon Publications, London.

Dennis, R.A. & Colfer, C.P. (2006) Impacts of land use and fire on the loss and degradation of lowland forest in 1983-2000 in East Kutai District, East Kalimantan, Indonesia. *Singapore Journal of Tropical Geography*, **27**, 30-48

Diniz-Filho, J.A.F., Hawkins, B.A., Mauricio Bini, L., De Marco Jr, P. & Blackburn, T.M. (2007) Are spatial regression methods a panacea or a Pandora's box? A reply to Beale et al. (2007). *Ecography*, **30**, 848-851

Dolédec, S., Chessel, D. & Gimaret-Carpentier, C. (2000) Niche separation in community analysis: A new method. *Ecology*, **81**, 2914–2927

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609-628

Duda, R.O., Hart, P.E. & Stork, D.G. (2001) *Pattern classification (2nd Edition)*. Wiley-interscience, New York, NY.

Dufrêne, M. & Legendre, P. (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, **67**, 345-366

Dunselman, J. (1939) Aantekeningen over zandpadangs in Borneo's Westerafdeeling. *De Tropische Natuur*, **28**, 71-82

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.C., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151

Endert, F.H. (1933) De proefbaanmetingen in Sangkoelirang (Z.O. Borneo). *Tectona*, **26**, 50-92

FAO (2002) Terrastat; Global land resources GIS models and databases for poverty and food insecurity mapping. Land and Water Digital Media Series #20.

FAO (2006) *Global Forest Resource Assessment 2005*. FAO, Rome, Italy.

Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393-404

Ferrier, S., Watson, G., Pearce, J. & Drielsma, M. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New SouthWales. I. Species-level modelling. *Biodiversity and Conservation*, **11**, 2275-2307

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49

Gathorne-Hardy, F.J., Syaukani, Davies, R.G., Eggleton, P. & Jones, D.T. (2002) Quaternary rainforest refugia in south-east Asia: using termites (Isoptera) as indicators. *Biological Journal of the Linnean Society*, **75**, 453-466

GLC2000 (2003) Global Land Cover 2000 database. European Commission, Joint Research Centre. http://www-gem.jrc.it/glc2000

Gotelli, N.J. & McGill, B.J. (2006) Null versus neutral models: What's the difference? *Ecography*, **29**, 793-800

Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Peterson, A.T. & Loiselle, B.A. (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239-247

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, **19**, 497-503

Graham, C.H., Moritz, C. & Williams, S.E. (2006) Habitat history improves prediction of biodiversity in rainforest fauna. *PNAS*, **103**, 632-636

Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, 2809–2815

Grytnes, J.A. & Beaman, J.H. (2006) Elevational species richness patterns for vascular plants on Mount Kinabalu, Borneo. *Journal of Biogeography*, **33**, 1838-1849

Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R. & Hastie, T. (2006) Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, **43**, 386-392

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147-186

Guralnick, R.P., Hill, A.W. & Lane, M. (2007) Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters*, **10**, 663-672

Hall, R. (1997) Cenozoic plate tectonic reconstructions of SE Asia. *Petroleum Geology of Southeast Asia* (ed. by A.J. Fraser, S.J. Matthews and R.W. Murphy), pp. 11-23. The Geological Society, London.

Hall, R. (1998) The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. *Biogeography and Geological Evolution of SE Asia* (ed. by R. Hall and J.D. Holloway), pp. 99-131. Backhuys Publishers, Leiden, The Netherlands.

Hall, R. (2009) Southeast Asia's changing palaeogeography. Blumea 54(1)

Hall, R., van Hattum, M.W.A. & Spakman, W. (2008) Impact of India-Asia collision on SE Asia: The record in Borneo. *Tectonophysics*, **451**, 366-389

Hallier, H. (1895) Rapport over de botanische tochten in Borneo's Westerafdeling gedurende de Borneo-expeditie 1893-94. *Natuurkundig Tijdschrift*, **54**, 406-449

Hannibal, L.W. (1950) Vegetation map of Indonesia. Planning Department, Forest Service, Jakarta, Indonesia

Hawkins, B.A., Diniz-Filho, J.A.F., Mauricio Bini, L., De Marco, P. & Blackburn, T.M. (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, **30**, 375-384

Heaney, L.R. (1991) A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change*, **19**, 53-61

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global

land areas. *International Journal of Climatology*, **25**, 1965-1978

Hirzel, A.H., Hauser, J., Chessel, D. & Perrini, N. (2002) Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data. *Ecology*, **83**, 2027-2036

Hortal, J., Garcia-Pereira, P. & Garcia-Barros, E. (2004) Butterfly species richness in mainland Portugal: predictive models of geographic distribution patterns. *Ecography*, **27**, 68-82

Hortal, J., Lobo, J.M. & Jiménez-Valverde, A. (2007) Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife, Canary Islands. *Conservation Biology*, **21**, 853-863

Jetz, W., Rahbek, C. & Colwell, R.K. (2004) The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, **7**, 1180-1191

Jiménez-Valverde, A. & Lobo, J.M. (2007) Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecologica*, **31**, 361-369

Johns, R.J. (1995) Malesia - An introduction. *Curtis's Botanical Magazine*, **12**, 52-62

Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelop models. *Ecological Applications*, **13**, 853-867

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401-413

Kennedy, M. (2000) *Understanding map projections*. Environmental Systems Research Institute, Inc., Redlands, USA.

Kier, G. & Barthlott, W. (2001) Measuring and mapping endemism and species richness: a new methodological approach and its application on the flora of Africa. *Biodiversity and Conservation*, **10**, 1513-1529

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T.H., Kuper, W., Kreft, H. & Barthlott, W. (2005) Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, **32**, 1107-1116

Küper, W., Sommer, J.H., Lovett, J.C. & Barthlott, W. (2006) Deficiency in African plant distribution data - missing pieces of the puzzle. *Botanical Journal of the Linnean Society*, **150**, 355-368

Langner, A., Miettinen, J. & Siegert, F. (2007) Land cover change 2002-2005 in Borneo and the role of fire derived from MODIS imagery. *Global Change Biology*, **13**, 2329–2340

Legendre, P. (2008) Studying beta diversity: ecological variation partitioning by multiple regression and canonical analysis. *Journal of Plant Ecology*, **1**, 3-8

Leigh Jr., E.G. (2004) How Wet are the Wet Tropics? *Tropical Forest Diversity and Dynamism* (ed. by E.C. Losos and E.G. Leigh Jr.), pp. 43-55. The University of Chicago Press, London.

Lennon, J.J., Koleff P., Greenwood J.J., Gaston K.J. 2004. Contribution of rarity and commonness to patterns of species richness. *Ecology Letters*, 7, 81-87.

Linder, H.P., Lovett, J.C., Mutke, J., Barthlott, W., Jürgens, N., Rebelo, T. & Küper, W. (2005) A numerical re-evaluation of the sub-Saharan phytochoria of mainland Africa. *Biologiske Skrifter*, **55**, 229-252

Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393

Lobo, J.M. & Martin-Piera, F. (2002) Searching for a Predictive Model for Species Richness of Iberian Dung Beetle Based on Spatial and Environmental Variables. *Conservation Biology*, **16**, 158-173

Loiselle, B.A., Jorgensen, P.M., Consiglio, T., Jimenez, I., Blake, J.G., Lohmann, L.G. & Montiel, O.M. (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography*, **35**, 105-116

MacKinnon, J. (1997) *Protected areas systems review of the Indo-Malayan realm*. ABC / WCMC / World Bank, Canterbury, England.

MacKinnon, K., Hata, G., Halim, H. & Mangalik, A. (1996) *The Ecology of Kalimantan*. Periplus Editions, Hong Kong.

Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921-931

McCune, B. & Grace, J.B. (2002) *Analysis of ecological communities*. MjM Software Design, Oregon.

McPherson, J.M. & Jetz, W. (2007) Effects of species' ecology on the accuracy of distribution models. *Ecography*, **30**, 135-151

McPherson, J.M., Jetz, W. & Rogers, D.J. (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, **41**, 811-823

Millennium-Ecosystem-Assessment (2005) *Ecosystems and Human Well-being: Biodiversity Synthesis*. World Resources Institute, Washington, DC.

Moerman, D.E. & Estabrook, G.F. (2006) The botanist effect: counties with maximal species richness tend to be home to universities and botanists. *Journal of Biogeography*, **33**, 1969-1974

Molengraaff, G.A.F. (1900) *Borneo-Expeditie - Geologische verkenningstochten in Centraal-Borneo (1893-94) - Atlas 22 bladen*. E.J. Brill & H. Gerlings, Leiden-Amsterdam.

Müller, S. (1857) *Reizen en onderzoekingen in den Indische Archipel gedaan op last der Nederlandsche Indische regering tusschen de jaren 1828 en 1836*. Frederik Muller, Amsterdam.

Mutke, J. & Barthlott, W. (2005) Patterns of vascular plant diversity at continental to global scales. *Biologiske Skrifter*, **55**, 521-531

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858

Newbery, D.M.C. (1991) Floristic variation within kerangas (heath) forest: re-evaluation of data from Sarawak and Brunei. *Plant Ecology*, **96**, 43-86

Nieuwenhuis, A.W. (1910) Tochten naar Boven-Boeloengan en de Apo-Kajan. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **27**, 263-306

Olden, J.D., Jackson, D.A. & Peres-Neto, P.R. (2002) Predictive Models of Fish Species Distributions: A Note on Proper Validation and Chance Predictions. *Transactions of the American Fisheries Society*, **131**, 329–336

Olson, D.M. & Dinerstein, E. (2002) The Global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical Garden*, **89**, 199-224

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the world: A new map of life on earth. *Bioscience*, **51**, 933-938

Parnell, J.A.N., Simpson, D.A., Moat, J., Kirkup, D.W., Chantaranothai, P., Boyce, P.C., Bygrave, P., Dransfield, S., Jebb, M.H.P., Macklin, J., Meade, C., Middleton, D.J., Muasya, A.M., Prajaksood, A., Pendry, C.A., Pooma, R., Suddee, S. & Wilkin, P. (2003) Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand. *Journal of*

Biogeography, **30**, 193-209
Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225-245
Pearce, J.L. & Boyce, M.S. (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, **43**, 405-412
Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102-117
Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704-1711
Perrin, P.M., Martin, J.R., Barron, S.J. & Roche, J.R. (2006) A cluster analysis approach to classifying Irish native woodlands. *Biology and Environment*, **106**, 261-275
Peterson, A.T. (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, **3**, 59-72
Peterson, A.T., Papes, M. & Eaton, M. (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, **30**, 550-560
Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259
Phillips, S.J., Dudík, M. & Schapire, R.E. (2004) A Maximum Entropy Approach to Species Distribution Modeling. In: *Twenty-First International Conference on Machine Learning, 2004*, p. 9
Potts, M.D., Ashton, P.S., Kaufman, L.S. & Plotkin, J.B. (2002) Habitat patterns in tropical rain forests: a comparison of 105 plots in northwest Borneo. *Ecology*, **83**, 2782-2797
Proches, S. (2005) The world's biogeographical regions: cluster analyses based on bat distributions. *Journal of Biogeography*, **32**, 607-614
Proctor, J., Anderson, J.M., Chai, P. & Vallack, H.W. (1983) Ecological Studies in Four Contrasting Lowland Rain Forests in Gunung Mulu National Park, Sarawak: I. Forest Environment, Structure and Floristics. *The Journal of Ecology*, **71**, 237-260
Proctor, J., Lee, Y.F., Langley, A.M., Munro, W.R.C. & Nelson, T. (1988) Ecological Studies on Gunung Silam, A Small Ultrabasic Mountain in Sabah, Malaysia. I. Environment, Forest Structure and Floristics. *The Journal of Ecology*, **76**, 320-340
Raes, N., Mols, J.B., Willemse, L.P.M. & Smets, E.F. (2009 - Chapter 2) Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images. *Blumea,* **54(1)**
Raes, N., Roos, M.C., slik, J.W.F., van Loon, E.E. & ter Steege, H. (accepted - Chapter 4) Botanical richness and endemicity patterns of Borneo derived from species distribution models. *Ecography*
Raes, N. & ter Steege, H. (2007 - Chapter 3) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727-736
Raes, N., van Loon, E.E., Roos, M.C. & ter Steege, H. (submitted - Chapter 5) The floristic regions of Borneo inferred from species distribution models.
Rahbek, C., Gotelli, N.J., Colwell, R.K., Entsminger, G.L., Rangel, T.F.L.V.B. & Graves, G.R. (2007) Predicting continental-scale patterns of bird species richness with spatially explicit models. *Proceedings of the Royal Society: Biological Sciences*, **274**, 165-174
Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Bini, L.M. (2006) Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography*, **15**, 321-327
Reddy, S. & Davalos, L.M. (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, **30**, 1719-1727
Reid, W.V. (1998) Biodiversity hotspots. *Trends in Ecology & Evolution*, **13**, 275-280
Romo, H., Garcia-Barros, E. & Lobo, J.M. (2006) Identifying recorder-induced geographic bias in an Iberian butterfly database. *Ecography*, **29**, 873-885
Roos, M.C., Keßler, P.J.A., Gradstein, R.S. & Baas, P. (2004) Species diversity and endemism of five major Malesian islands: diversity-area relationships. *Journal of Biogeography*, **31**, 1893-1908
Sakai, S. (2001) Phenological diversity in tropical forests. *Population Ecology*, **43**, 77-86
Schmidt, M., Kreft, H., Thiombiano, A. & Zizka, G. (2005) Herbarium collections and field data-based plant diversity maps for Burkina Faso. *Diversity and Distributions*, **11**, 509-516
Slatyer, C., Rosauer, D. & Lemckert, F. (2007) An assessment of endemism and species richness patterns in the Australian Anura. *Journal of Biogeography*, **34**, 583-596
Slik, J.W.F. (2004) El Niño droughts and their effects on tree species composition and diversity in tropical rain forests. *Oecologia*, **141**, 114–120
Slik, J.W.F., Poulsen, A.D., Ashton, P.S., Cannon, C.H., Eichhorn, K.A.O., Kartawinata, K., Lanniari, I., Nagamasu, H., Nakagawa, M., van Nieuwstadt, M.G.L., Payne, J., Saridan, A., Sidiyasa, K., Verburg, R.W., Webb, C.O. & Wilkie, P. (2003) A floristic analysis of the lowland dipterocarp forests of Borneo. *Journal of Biogeography*, **30**, 1517-1531
Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**, 1-10
Sodhi, N.S., Koh, L.P., Brook, B.W. & Ng, P.K.L. (2004) Southeast Asian biodiversity: an impending disaster. *Trends in Ecology & Evolution*, **19**, 654-660
Soethe, N., Lehmann, J. & Engels, C. (2008) Nutrient availability at different altitudes in a tropical montane forest in Ecuador. *Journal of Tropical Ecology*, **24**, 397-406
Stibig, H.J., Belward, A.S., Roy, P.S., Rosalina-Wasrin, U., Agrawal, S., Joshi, P.K., Hildanus, Beuchle, R., Fritz, S., Mubareka, S. & Giri, C. (2007) A land-cover map for South and Southeast Asia derived from SPOT-VEGETATION data. *Journal of Biogeography*, **34**, 625-637
Stockwell, D. & Peters, D. (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, **13**, 143-158
Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, **148**, 1–13
Stolk, J.J. (1907) Opsporing van den zwervenden stam der Penjaboeng-Poenan's op de waterscheiding der Barito met de Mahakan en Kapoeas (Midden-Borneo), in Oct. 1905. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **24**, 1-27
Stone, R. (2007) Last-Gasp Effort to Save Borneo's Tropical Rainforests. *Science*, **317**, 192
Svenning, J.-C., Normand, S. & Skov, F. (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, **31**, 316-326
Swets, J.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285-1293

Swets, J.A., Dawes, R.M. & Monahan, J. (2000) Better decisions through science. *Scientific American*, **283**, 82-87

ter Steege, H., Pitman, N., Sabatier, D., Castellanos, H., van der Hout, P., Daly, D.C., Silveira, M., Phillips, O., Vasquez, R., van Andel, T., Duivenvoorden, J., de Oliveira, A.A., Ek, R., Lilwah, R., Thomas, R., van Essen, J., Baider, C., Maas, P., Mori, S., Terborgh, J., Vargas, P.N., Mogollon, H. & Morawetz, W. (2003) A spatial model of tree alpha-diversity and tree density for the Amazon. *Biodiversity and Conservation*, **12**, 2255-2277

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234-2250

Urban, D., Goslee, S., Pierce, K. & Lookingbill, T. (2002) Extending community ecology to landscapes. *Ecoscience*, **9**, 200-212

van Balen, S. (1996) The ornithological importance of the Danau Sentarum Wildlife reserve in West Kalimantan. *Sarawak Museum Journal*, **50**, 123-144

van Balgooy, M.M.J., Hovenkamp, P.H. & van Welzen, P.C. (1996) Phytogeography of the Pacific - floristic and historical distribution patterns in plants. *The origin and evolution of Pacific Island biotas, New Guinea to Eastern Polynesia: patterns and processes* (ed. by A. Keast and S.E. Miller), pp. 191-213. SPB Academic Publishing, Amsterdam, The Netherlands.

van Diessen, J.R., van den Brink, P. & Bek, W. (2004) *Grote Atlas van Nederlands Oost-Indië / Comprehensive atlas of the Netherlands East Indies*. Asia Maior Publishers/KNAG, second extended edition, Zierikzee.

van Steenis, C.G.G.J. (1935a) Maleische vegetatieschetsen I; Toelichting bij de plantengeografische kaart van Nederlandsch Oost Indië. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **52**, 25-67

van Steenis, C.G.G.J. (1935b) Maleische vegetatieschetsen II. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **52**, 171-203; 365-398

van Steenis, C.G.G.J. (1948) Hoofdlijnen van de plantengeographie van de Indische archipel op grond van de verspreiding de Phanerogamen geslachten. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **65**, 193-208

van Steenis, C.G.G.J. (1950) The delimitation of Malaysia and its main plant geographical divisions. *Flora Malesiana series 1* (ed. by C.G.G.J. van Steenis), pp. LXX-LXXV. Noordhoff-Kolff N.V., Djakarta.

van Steenis, C.G.G.J. (1958a) Concise history of the map and an outline of Malaysian vegetation mapping and literature. *Proceedings of the Symposium on Humid Tropics Vegetation, Kandy*. Unesco, Paris.

van Steenis, C.G.G.J. (1958b) Vegetation map of Malaysia (in collaboration with UNESCO for the UNESCO Humid Tropics Research Project).

van Steenis, C.G.G.J. (1979) Plant-geography of east Malesia. *Botanical Journal of the Linnean Society*, **79**, 97-178

van Walcheren, E.W.F. (1907) Eene reis naar de bovenstreken van Boeloengan (Midden-Borneo) (12 Nov. 1905 - 11 April 1906). *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **24**, 755-844

van Welzen, P.C. & Slik, J.W.F. (2009) Patterns in species richness and composition of plant families in the Malay Archipelago. *Blumea,* **54(1)**

van Welzen, P.C., Slik, J.W.F. & Alahuhta, J. (2005) Plant distribution patterns and plate tectonics in Malesia. *Biologiske Skrifter*, **55**, 199-217

Verstappen, H.T. (1997) The effect of climatic change on southeast Asian geomorphology. *Journal of Quaternary Science*, **12**, 413-418

Voris, H.K. (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153-1167

Wakita, K. & Metcalfe, I. (2005) Ocean Plate Stratigraphy in East and Southeast Asia. *Journal of Asian Earth Sciences*, **24**, 679-702

Wallace, A.R. (1860) On the zoological geography of the Malay Archipelago. *Journal of the Linnean Society of London* **4**, 172-184

Walsh, R.P.D. (1996) Climate. *The Tropical Rain Forest* (ed. by P.W. Richards), pp. 159-205. Cambridge University Press, Cambridge, U.K.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyári, Á.S., Perkins, S.L. & Guralnick, R.P. (2007) Locating Pleistocene Refugia: Comparing Phylogeographic and Ecological Niche Model Predictions. *PLoS ONE*, **2**, e563

WDPA (2007) World Database on Protected Areas, web-download. UNEP-WCMC and IUCN World Commission on Protected Areas December 2007. http://www.unep-wcmc.org/wdpa/

Whitmore, T.C. (1984a) *Tropical rain forests of the Far East*. Clarendon Press, Oxford.

Whitmore, T.C. (1984b) A vegetation map of Malesia at scale 1:5 million. *Journal of Biogeography*, **11**, 461-471

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M. & Willis, K.J. (2005) Conservation Biogeography: assessment and prospect. *Diversity and Distributions*, **11**, 3-23

Wikramanayake, E., Dinerstein, E., Loucks, C.J., Olson, D.M., Morrison, J. & Lamoureux, J. (2002) *Terrestrial ecoregions of the Indo-Pacific: A conservation assessment*. Island Press, Washington, D.C.

Winkler, H. (1927) *Die planzendecke Süd-Ost Borneos?* Institüt für algemeine Botanik, Hamburg.

Wisz, M.S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14, 763-773.

Witkamp, H. (1932) Langs de Mahakam. *Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap*, **49**, 30-65

Witt, C.C. & Maliakal-Witt, S. (2007) Why are diversity and endemism linked on islands? *Ecography*, **30**, 331-333

WWF-Germany (2005) *Borneo: Treasure Island at Risk*. WWF-Germany, Frankfurt am Main.

WWF & IUCN (1995) *Centres of Plant Diversity, a guide and strategy for their conservation; Volume 2 Asia, Australasia and the Pacific*. The world Wide Fund for Nature (WWF) and the IUCN - The World Conservation Union.

Yesson, C. & Culham, A. (2006a) Phyloclimatic Modeling: Combining Phylogenetics and Bioclimatic Modeling. *Systematic Biology*, **55**, 785 - 802

Yesson, C. & Culham, A. (2006b) A phyloclimatic study of Cyclamen. *BMC Evolutionary Biology*, **6**, 1-23

Zaniewski, A.E., Lehmann, A. & Overton, J.M. (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, **157**, 261-280

Zollinger, H. (1857) Over het begrip en den omvang eener Flora Malesiana. *Natuurkundig Tijdschrift voor Nederlandsch-Indië*, **13**, 293-322

Figure S4.1

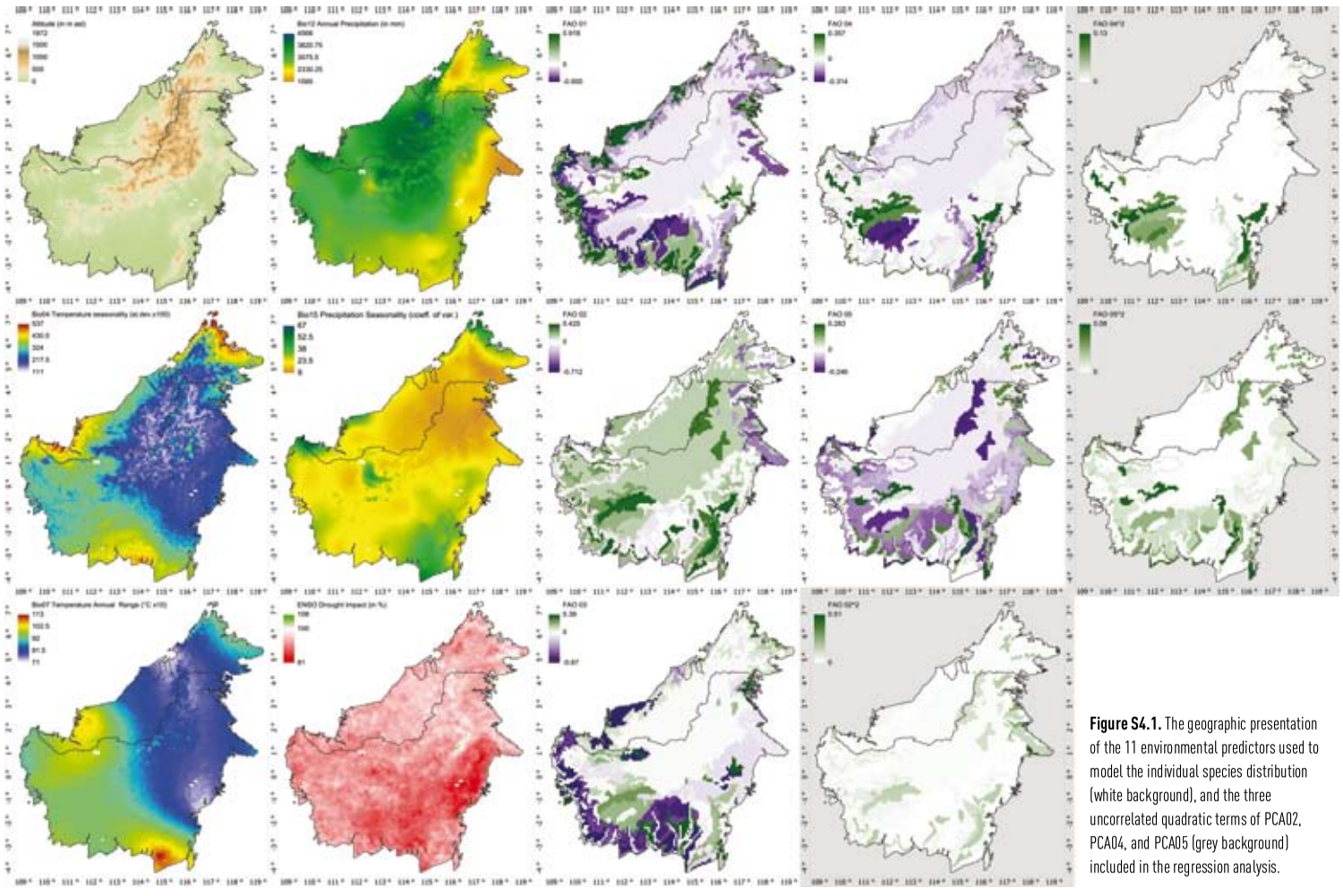


Figure S4.1. The geographic presentation of the 11 environmental predictors used to model the individual species distribution (white background), and the three uncorrelated quadratic terms of PCA02, PCA04, and PCA05 (grey background) included in the regression analysis.

Figure S4.2

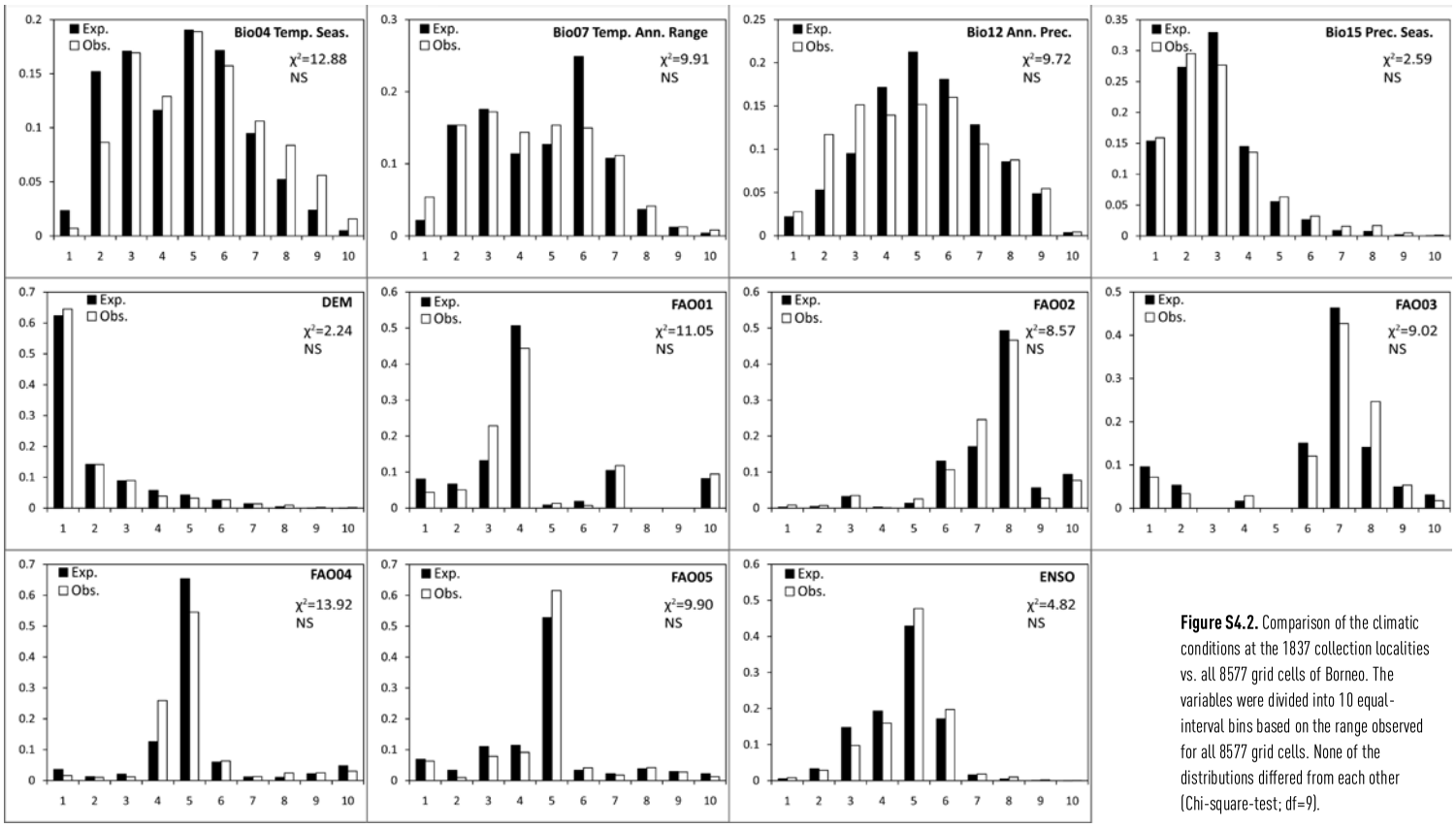


Figure S4.2. Comparison of the climatic conditions at the 1837 collection localities vs. all 8577 grid cells of Borneo. The variables were divided into 10 equal-interval bins based on the range observed for all 8577 grid cells. None of the distributions differed from each other (Chi-square-test; df=9).

Figure S4.3

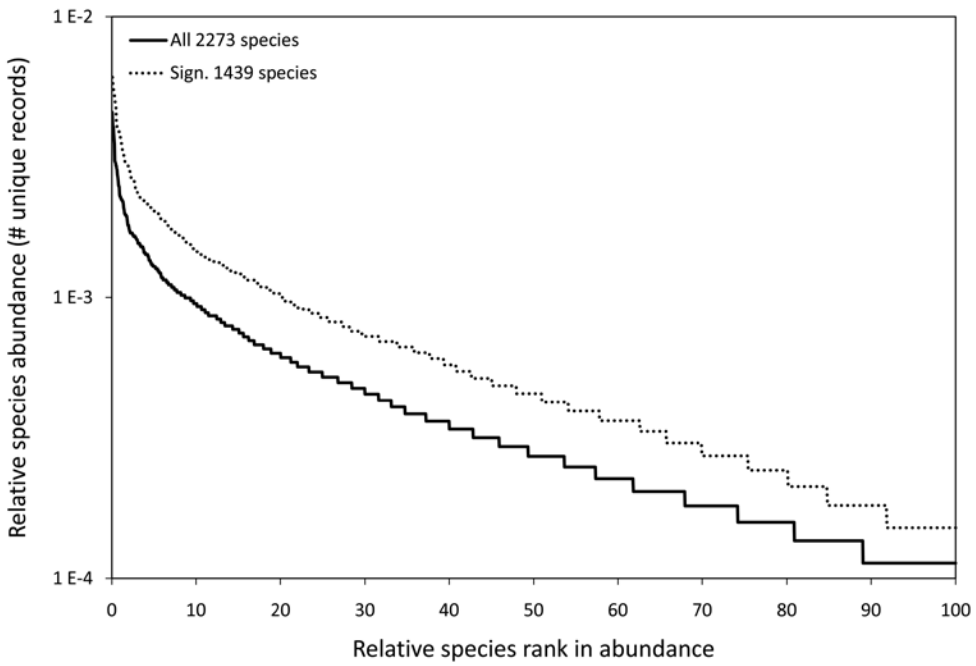


Figure S4.3. Relative species rank abundance curves based on the number of unique records per species for all 2273 modelled species, and for the 1439 species with a significant SDM.

Figure S5.1

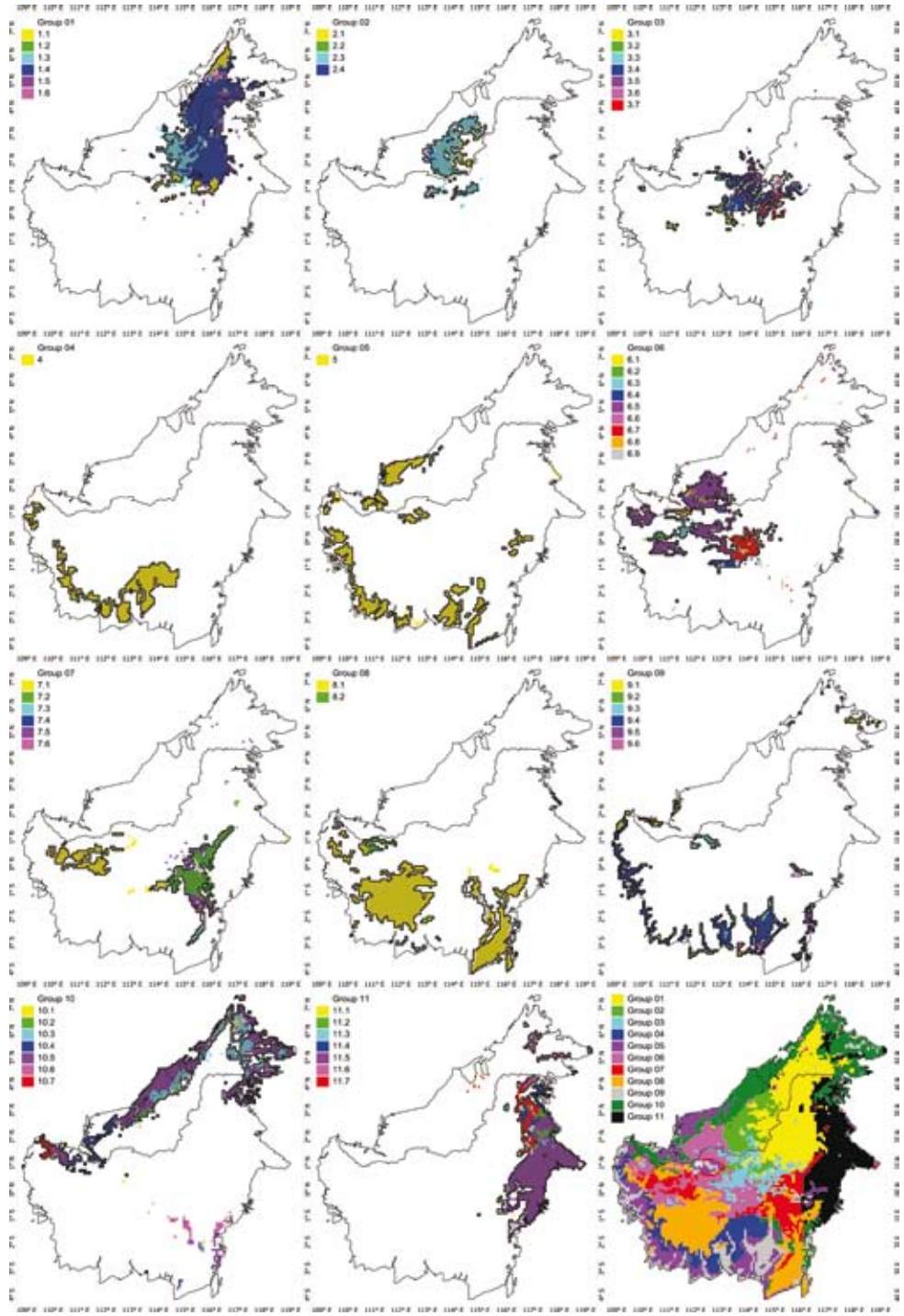


Figure S5.1. Geographical presentation of the terminal nodes of the CART per floristic region indicated by Group 01-11. Hatched areas indicate the geographical extent of the corresponding floristic regions derived from the species presence/absence cluster analysis. Lower right figure shows the CART areas belonging to each group.

Figure S5.2

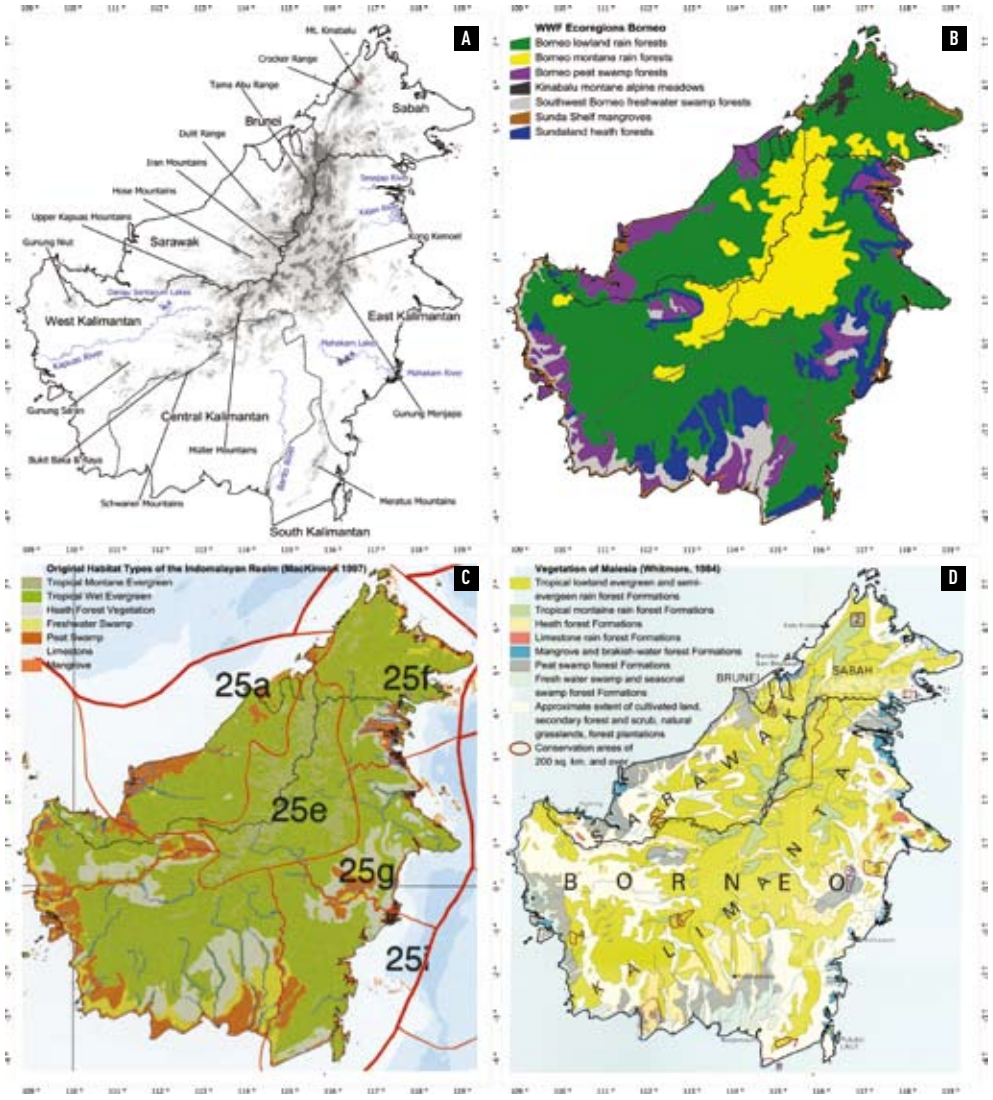


Figure S5.2. A) Map indicating the Bornean administrative provinces of Malaysia and Indonesia, the country of Brunei, the major mountain ranges and large rivers. Areas between 500-1000m in light-grey; areas above 1000m in dark-grey; Mt. Kinabalu in red (≥ 2600m) (derived from the SRTM 90m resolution altitude data (<http://srtm.csi.cgiar.org>)). **B)** WWF ecoregions of Borneo (<http://www.worldwildlife.org/science/data/item1875.html>; downloaded 06/02/2008); colours similar to floristic regions map. **C)** Detail from the 'Original Habitat Types of the Indomalayan Realm' of MacKinnon (1997). **D)** Detail from the 'Vegetation of Malesia' map of Whitmore (1984). Note that the Limestone areas of western Borneo identified by MacKinnon (C) are erroneously mapped by the WWF (B) as heath forest. All maps in geographic projection.

Table S3.1. Results for the 116 Shorea distribution models. The number of unique grid cells in which a species was collected is indicated by '# Records'. The SDM AUC values are given by 'AUC'. Bold values are significantly deviating from both null-models, italic only from the random null-model (p<0.05). Fitted 95% confidence interval (C.I.) AUC values for the random null-model are given by '95% C.I. All'. Fitted 95% C.I. AUC values for the biased corrected null-model are given by '95% C.I. Bias'. Asterisks indicate significant deviance of the SDM from the null-models (p<0.05). The 'Area (in %)' indicates the percentage of the total area of Borneo for which a species was predicted to be present, after the sensitivity-specificity sum maximization threshold was applied to the continuous probabilistic Maxent predictions (only given for significant SDMs).

#	Species	# Records	AUC	95% C.I. All	95% C.I. Bias	Area (in %)
1	Shorea bullata P.S.Ashton	5	**0.9878**	0.9498*	0.9625*	3.8
2	S. carapae Ashton	5	0.9014	0.9498	0.9625	
3	S. flemmichii Symington	5	0.8854	0.9498	0.9625	
4	S. hemsleyana King ex Foxw. subsp. grandiflora Brandis P.S.Ashton	5	**0.9824**	0.9498*	0.9625*	4.8
5	S. kudatensis G.H.S.Wood ex Meijer	5	**0.9830**	0.9498*	0.9625*	6.4
6	S. parvistipulata Heim subsp. nebulosa Meijer Ashton	5	**0.9868**	0.9498*	0.9625*	2.4
7	S. pubistyla P.S.Ashton	5	**0.9910**	0.9498*	0.9625*	1.8
8	S. revoluta P.S.Ashton	5	**0.9976**	0.9498*	0.9625*	1.2
9	S. elliptica Burck	6	**0.9892**	0.9324*	0.9473*	5.4
10	S. obovoidea Slooten	6	**0.9935**	0.9324*	0.9473*	1.7
11	S. parvistipulata Heim subsp. albifolia P.S.Ashton	6	0.8618	0.9324	0.9473	
12	S. rubella P.S.Ashton	6	**0.9663**	0.9324*	0.9473*	7.7
13	S. symingtonii G.H.S.Wood	6	**0.9898**	0.9324*	0.9473*	2.2
14	S. agami P.S.Ashton subsp. diminuta Ashton	7	0.9050	0.9150	0.9321	
15	S. asahi P.S.Ashton	7	**0.9411**	0.9150*	0.9321*	17.0
16	S. biawak P.S.Ashton	7	**0.9776**	0.9150*	0.9321*	8.1
17	S. iliasii P.S.Ashton	7	**0.9516**	0.9150*	0.9321*	4.1
18	S. laxa Slooten	7	**0.9753**	0.9150*	0.9321*	6.4
19	S. macrobalanos P.S.Ashton	7	0.8777	0.9150	0.9321	
20	S. pachyphylla Ridl. ex Symington	7	**0.9671**	0.9150*	0.9321*	13.6
21	S. uliginosa Foxw.	7	**0.9376**	0.9150*	0.9321*	14.8
22	S. macrantha Brandis	8	*0.9116*	0.8977*	0.9169	4.4
23	S. agami P.S.Ashton subsp. agami P.S.Ashton	9	**0.9269**	0.8803*	0.9017*	20.1
24	S. curtisii Dyer ex King	9	**0.9296**	0.8803*	0.9017*	18.5
25	S. cuspidata P.S.Ashton	9	**0.9884**	0.8803*	0.9017*	5.0
26	S. falcifera Dyer	9	**0.9509**	0.8803*	0.9017*	25.0
27	S. mujongensis P.S.Ashton	9	*0.8916*	0.8803*	0.9017	7.7
28	S. ovalis Korth. Blume subsp. sarawakensis P.S.Ashton	9	**0.9580**	0.8803*	0.9017*	15.3
29	S. sagittata Ashton	9	**0.9492**	0.8803*	0.9017*	16.4
30	S. teysmanniana Dyer	9	**0.9374**	0.8803*	0.9017*	15.8
31	S. almon Foxw.	10	*0.9196*	0.8900*	0.9201	6.7
32	S. dasyphylla Foxw.	10	0.8769	0.8900	0.9201	
33	S. gratissima Wall. ex Kurz Dyer	10	**0.9916**	0.8900*	0.9201*	3.8
34	S. inappendiculata Burck	10	**0.9267**	0.8900*	0.9201*	9.8
35	S. myrionerva Symington	10	**0.9292**	0.8900*	0.9201*	25.1
36	S. parvistipulata Heim subsp. parvistipulata	10	0.8693	0.8900	0.9201	
37	S. slootenii G.H.S.Wood	10	*0.8990*	0.8900*	0.9201	12.0
38	S. albida Symington	11	**0.9746**	0.8821*	0.9097*	3.5
39	S. angustifolia Ashton	11	*0.8842*	0.8821*	0.9097	30.0
40	S. crassa P.S.Ashton	11	**0.9439**	0.8821*	0.9097*	19.0
41	S. foxworthyi Symington	11	**0.9555**	0.8821*	0.9097*	17.7
42	S. longiflora Brandis Symington	11	0.8696	0.8821	0.9097	
43	S. acuminatissima Symington	12	**0.9737**	0.8741*	0.9021*	12.2
44	S. faguetioides P.S.Ashton	12	*0.8799*	0.8741*	0.9021	24.5
45	S. isoptera P.S.Ashton	12	**0.9808**	0.8741*	0.9021*	6.9

Table S3.1. (continued).

#	Species	# Records	AUC	95% C.I. All	95% C.I. Bias	Area (in %)
46	S. rugosa Heim	12	*0.8838*	0.8741*	0.9021	19.1
47	S. andulensis P.S.Ashton	13	**0.9623**	0.8661*	0.8973*	10.2
48	S. balangeran Korth. Burck	13	**0.9555**	0.8661*	0.8973*	13.9
49	S. platycarpa Heim	13	0.8557	0.8661	0.8973	
50	S. stenoptera Burck	13	**0.9295**	0.8661*	0.8973*	15.3
51	S. domatiosa P.S.Ashton	14	0.9701	0.8581*	0.8953*	14.6
52	S. macropterafolia P.S.Ashton	14	**0.9445**	0.8581*	0.8953*	13.1
53	S. splendida Ashton	14	*0.8736*	0.8581*	0.8953	38.0
54	S. flaviflora G.H.S.Wood ex P.S.Ashton	15	*0.9622*	0.9480*	0.9622	16.1
55	S. macroptera Dyer subsp. bailloni	15	**0.9737**	0.9480*	0.9622*	5.9
56	S. retusa Meijer	15	**0.9695**	0.9480*	0.9622*	6.4
57	S. confusa P.S.Ashton	16	**0.9882**	0.9438*	0.9595*	8.1
58	S. exelliptica Meijer	16	0.9569	0.9438*	0.9595	17.1
59	S. longisperma Roxb.	16	0.9416	0.9438	0.9595	
60	S. platyclados Slooten ex Endert	16	0.9541	0.9438*	0.9595	8.8
61	S. falciferoides Foxw.	17	**0.9702**	0.9397*	0.9569*	7.9
62	S. patoiensis P.S.Ashton	17	0.9501	0.9397*	0.9569	17.4
63	S. balanocarpoides Symington	18	**0.9829**	0.9357*	0.9543*	5.1
64	S. pilosa P.S.Ashton	18	0.9434	0.9357*	0.9543	6.6
65	S. superba Foxw.	18	**0.9846**	0.9357*	0.9543*	8.7
66	S. macroptera Dyer	19	0.9241	0.9318	0.9517	
67	S. ovalis Korth. Blume subsp. ovalis	19	**0.9561**	0.9318*	0.9517*	13.5
68	S. brunnescens P.S.Ashton	20	**0.9778**	0.9280*	0.9492*	5.3
69	S. hypoleuca Meijer	20	**0.9822**	0.9280*	0.9492*	8.5
70	S. monticola P.S.Ashton	20	**0.9822**	0.9280*	0.9492*	5.2
71	S. scrobiculata Burck	21	**0.9646**	0.9243*	0.9468*	8.6
72	S. virescens Parijs	21	**0.9550**	0.9243*	0.9468*	16.6
73	S. xanthophylla Symington	21	**0.9662**	0.9243*	0.9468*	6.9
74	S. coriacea Burck	22	**0.9643**	0.9206*	0.9444*	12.0
75	S. palembanica Miq.	22	0.9426	0.9206*	0.9444	6.2
76	S. venulosa G.H.S.Wood ex Meijer	22	**0.9818**	0.9206*	0.9444*	5.6
77	S. mecistopteryx Ridl.	23	0.9368	0.9170*	0.9420	13.5
78	S. obscura Meijer	24	0.9251	0.9135*	0.9397	8.9
79	S. macroptera Dyer subsp. sandakanensis Symington P.S.Ashton	25	**0.9575**	0.9101*	0.9374*	5.4
80	S. ovata Dyer ex Brandis	25	**0.9818**	0.9101*	0.9374*	6.7
81	S. argentifolia Symington	27	**0.9396**	0.9035*	0.9330*	17.8
82	S. havilandi Brandis	27	**0.9762**	0.9035*	0.9330*	10.5
83	S. maxwelliana King	27	**0.9604**	0.9035*	0.9330*	6.7
84	S. ochracea Symington	27	0.9203	0.9035*	0.9330	14.2
85	S. polyandra P.S.Ashton	27	**0.9434**	0.9035*	0.9330*	22.7
86	S. rubra P.S.Ashton	27	**0.9786**	0.9035*	0.9330*	7.3
87	S. agami P.S.Ashton	28	**0.9696**	0.9003*	0.9308*	11.3
88	S. lamellata Foxw.	28	**0.9525**	0.9003*	0.9308*	10.1
89	S. quadrinervis Slooten	28	**0.9515**	0.9003*	0.9308*	8.0
90	S. kunstleri King	30	0.9136	0.8942*	0.9267	18.0
91	S. hopeifolia Heim Symington	32	0.8958	0.8883*	0.9226	16.7
92	S. scaberrima Burck	32	**0.9464**	0.8883*	0.9226*	10.2
93	S. ferruginea Dyer ex Brandis	34	**0.9380**	0.8827*	0.9188*	22.1
94	S. scabrida Symington	34	**0.9494**	0.8827*	0.9188*	18.5
95	S. amplexicaulis Ashton	35	**0.9246**	0.8800*	0.9169*	11.0
96	S. beccariana Burck	35	0.9047	0.8800*	0.9169	25.7
97	S. parvifolia Dyer subsp. parvifolia	35	0.9104	0.8800*	0.9169	18.1
98	S. atrinervosa Symington	37	**0.9365**	0.8748*	0.9133*	9.1

Table S3.1. (continued).

#	Species	# Records	AUC	95% C.I. All	95% C.I. Bias	Area (in %)
99	S. ovalis Korth. Blume	37	**0.9546**	0.8748*	0.9133*	20.0
100	S. bracteolata Dyer	38	**0.9533**	0.8722*	0.9115*	17.4
101	S. guiso Blanco Blume	38	**0.9559**	0.8722*	0.9115*	17.7
102	S. parvifolia Dyer	38	**0.9243**	0.8722*	0.9115*	21.9
103	S. johorensis Foxw.	39	**0.9358**	0.8698*	0.9098*	20.6
104	S. faguetiana Heim	43	**0.9348**	0.8605*	0.9032*	7.4
105	S. macrophylla de Vriese P.S.Ashton	46	0.8819	0.8542*	0.8987	26.8
106	S. smithiana Symington	46	**0.9454**	0.8542*	0.8987*	10.5
107	S. fallax Meijer	47	**0.9404**	0.8522*	0.8972*	10.0
108	S. parvistipulata Heim	52	**0.9206**	0.8428*	0.8904*	12.8
109	S. seminis de Vriese Slooten	55	**0.8965**	0.8377*	0.8866*	30.2
110	S. gibbosa Bandis	68	**0.9042**	0.8191*	0.8730*	18.3
111	S. multiflora Burck Symington	68	**0.9272**	0.8191*	0.8730*	16.9
112	S. laevis Ridl.	69	**0.8778**	0.8178*	0.8721*	19.9
113	S. pauciflora King	73	0.8666	0.8131*	0.8687	12.5
114	S. pinanga Scheff.	73	0.8609	0.8131*	0.8687	22.1
115	S. leprosula Miq.	86	**0.8872**	0.7988*	0.8593*	23.5
116	S. parvifolia Dyer subsp. velutinata P.S.Ashton	92	0.8475	0.7923*	0.8557	21.2

Table S4.1. Pairwise Pearson's correlation coefficients (r) for the 14 environmental predictors. Legend: DEM - Digital Elevation Model; Bio04 – Temperature seasonality; Bio07 – Temperature annual range; Bio12 – Annual precipitation; Bio15 – Precipitation seasonality; PCA01-05 – PCA soil axes 1-5; ENSO – Drought impact by El Ninõ events on NDVI values.Table S5.1. Results of the Indicator Species Analysis (ISA) on 11 'Cluster Groups' and 8577 grid cells. Species are ranked on Indicator Values (IndVal) per cluster group. 'Area (in %)' gives both the percentages of the surface of Borneo covered by a species derived from its significant Species Distribution Model (SDM). Values in the table give the rounded 'percentage of perfect indication'.

	DEM	Bio04	Bio07	Bio12	Bio15	PCA01	PCA02	PCA03	PCA04	PCA05	ENSO	PCA02²	PCA04²
Bio04	-0.512												
Bio07	-0.361	**0.737**											
Bio12	0.207	-0.167	0.030										
Bio15	-0.412	0.514	0.507	-0.254									
PCA01	-0.216	0.194	0.142	-0.030	0.185								
PCA02	0.274	-0.078	0.094	0.214	0.040	0.138							
PCA03	0.226	-0.214	-0.202	0.066	-0.168	-0.306	-0.036						
PCA04	-0.064	0.026	0.081	-0.095	0.149	0.034	0.151	-0.018					
PCA05	-0.136	0.033	0.004	-0.062	0.048	0.068	-0.011	0.021	-0.010				
ENSO	0.372	-0.078	-0.099	0.274	-0.390	-0.053	-0.003	0.128	-0.083	0.034			
PCA02²	0.078	-0.035	-0.023	-0.208	-0.026	-0.028	0.044	0.022	0.364	0.230	0.084		
PCA04²	-0.085	0.134	0.193	-0.109	0.184	-0.230	0.323	0.285	0.464	0.119	-0.126	0.331	
PCA05²	-0.062	0.149	0.112	-0.140	0.094	-0.170	0.060	-0.078	0.049	0.112	-0.058	0.334	0.292

Table S5.1. Results of the Indicator Species Analysis (ISA) on 11 'Cluster Groups' and 8577 grid cells. Species are ranked on Indicator Values (IndVal) per cluster group. 'Area (in %)' gives both the percentages of the surface of Borneo covered by a species derived from its significant Species Distribution Model (SDM). Values in the table give the rounded 'percentage of perfect indication'.

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1239	Cunoniaceae	*Weinmannia aphanoneura* Airy Shaw	1	69	12.3	85	0	12	0	0	0	0	0	0	1	1
1324	Orchidaceae	*Appendicula cristata* Blume	1	65	12.85	82	0	5	0	0	0	1	0	0	7	5
1356	Orchidaceae	*Bulbophyllum rhizomatosum* Alb. & Schwein.	1	64.3	14.2	73	0	25	0	0	0	0	0	0	1	1
1364	Orchidaceae	*Chelonistele ingloria* J.J.Sm. Carr	1	64.3	12.94	78	6	13	0	0	0	0	0	0	1	2
582	Fagaceae	*Quercus valdinervosa* Soepadmo	1	61.4	13.64	75	0	15	0	0	0	0	0	0	8	1
658	Symplocaceae	*Symplocos ophirensis* C.B.Clarke subsp. *cumingiana* Brand Noot. var. *cumingiana*	1	61.2	10.05	92	0	1	0	0	0	0	0	0	5	2
1204	Myristicaceae	*Horsfieldia subalpina* J.Sinclair subsp. *kinabaluensis* W.J.de Wilde	1	60.6	16.65	64	1	31	0	0	0	0	0	0	2	3
1168	Loranthaceae	*Helixanthera spicata* Danser	1	60	15.73	69	6	9	0	0	0	2	5	0	6	5
1240	Cunoniaceae	*Weinmannia fraxinea* D.Don Miq.	1	59.6	13.84	74	2	11	0	0	0	3	4	1	3	3
461	Ericaceae	*Rhododendron himantodes* Sleumer	1	58.9	14.41	71	1	17	0	0	0	0	4	0	3	3
650	Symplocaceae	*Symplocos laeteviridis* Stapf var. *basirotunda* Noot.	1	57	13.83	74	3	6	0	0	0	0	5	0	9	4
1433	Orchidaceae	*Thelasis variabilis* Ames & C.Schweinf.	1	56.9	13.7	71	0	20	0	0	0	0	4	0	2	2
1381	Orchidaceae	*Coelogyne rhadobulbon* Schltr.	1	56.2	19.73	57	4	22	0	0	0	4	0	0	4	9
881	Araliaceae	*Schefflera trineura* Frodin	1	56.1	17.07	60	10	23	0	0	0	2	0	0	3	3
492	Juglandaceae	*Engelhardtia roxburghiana* Wall.	1	54.1	17.58	62	3	10	0	0	0	0	1	0	12	11
1389	Orchidaceae	*Dendrobium corallorhizon* J.J.Sm.	1	53.9	16.96	64	0	6	0	0	0	1	3	0	5	20
75	Euphorbiaceae	*Antidesma neurocarpum* Miq. var. *neurocarpum*	1	52.9	21.76	54	6	11	0	0	0	0	0	0	13	16
998	Rosaceae	*Prunus arborea* Blume Kalkman var. *alticola* Kalkman	1	52.7	17.27	58	7	25	0	0	0	0	1	0	1	8
1361	Orchidaceae	*Calanthe pulchra* Blume Lindl.	1	52.5	14.53	65	1	25	0	0	0	0	4	0	2	3
458	Ericaceae	*Rhododendron crassifolium* Stapf	1	51.4	18.08	54	10	30	0	0	0	0	0	0	5	0
1358	Orchidaceae	*Bulbophyllum striatellum* Ridl.	1	51.2	16.85	56	13	26	0	0	0	0	0	0	2	2
1341	Orchidaceae	*Bulbophyllum ecornutum* J.J.Sm. J.J.Sm.	1	50.6	11.89	74	4	4	0	0	0	1	0	0	6	10
1338	Orchidaceae	*Bulbophyllum coniferum* Ridl.	1	49.3	15.1	61	0	30	0	0	0	0	0	0	9	0
1375	Orchidaceae	*Coelogyne longibulbosa* Ames & C.Schweinf.	1	49.3	10.99	73	0	26	0	0	0	0	0	0	0	1
1345	Orchidaceae	*Bulbophyllum membranifolium* Hook.f.	1	49	14.54	62	0	29	0	0	0	0	3	0	0	6
1236	Nepenthaceae	*Nepenthes lowii* Hook.f.	1	48.7	11.04	74	0	19	0	0	0	0	0	0	2	6
1376	Orchidaceae	*Coelogyne moultonii* J.J.Sm.	1	48.7	17.3	57	1	25	0	0	0	0	11	0	2	3
1157	Illiciaceae	*Illicium stapfii* Merr.	1	48.1	15.34	59	0	32	0	0	0	0	3	0	5	1
1192	Myristicaceae	*Horsfieldia endertii* W.J.de Wilde	1	47.6	7.19	96	0	1	0	0	0	0	0	0	1	1
514	Clethraceae	*Clethra longispicata* J.J.Sm.	1	47.5	7.92	91	0	4	0	0	0	0	2	0	1	3
1399	Orchidaceae	*Dendrochilum imbricatum* Ames	1	46.9	8.23	89	0	0	0	0	0	0	0	0	10	1
490	Ericaceae	*Vaccinium uroglossum* Sleumer	1	46.4	13.91	60	0	36	0	0	0	0	2	0	0	1
654	Symplocaceae	*Symplocos laeteviridis* Stapf var. *pauciflora* Noot.	1	46.4	13.49	62	10	21	0	0	0	0	1	0	3	2
85	Euphorbiaceae	*Aporosa chondroneura* Airy Shaw Schot	1	46.1	20.89	47	20	24	0	0	0	0	0	0	4	5
891	Aristolochiaceae	*Aristolochia minutiflora* Ridl. ex Gamble	1	46.1	15.55	58	0	27	0	0	0	1	3	0	4	6
1004	Rosaceae	*Prunus oocarpa Stapf* Kalkman	1	45.6	11.37	69	0	25	0	0	0	0	1	0	0	5
488	Ericaceae	*Vaccinium retivenium* Sleumer	1	45.3	17.14	55	0	27	0	0	1	1	0	0	13	3
1420	Orchidaceae	*Nabaluia angustifolia* de Vogel	1	45	7.18	92	0	6	0	0	0	0	0	0	0	1
1424	Orchidaceae	*Pholidota clemensii* Ames	1	43.6	8.46	80	0	19	0	0	0	0	0	0	1	0
486	Ericaceae	*Vaccinium pachydermum* Stapf	1	43.3	13.78	62	12	4	0	0	0	0	0	0	16	5
1178	Loranthaceae	*Macrosolen melintangensis* Korth. Miq.	1	43.3	11.85	69	0	4	0	6	0	0	0	0	12	9
1347	Orchidaceae	*Bulbophyllum minutulum* Ridl.	1	43	12.57	66	0	14	0	0	0	0	8	0	1	11
466	Ericaceae	*Rhododendron micromalayanum* Sleumer	1	42.9	8.02	86	2	1	0	0	0	0	0	0	3	8
473	Ericaceae	*Rhododendron quadrasianum* S.Vidal var. *villosum* J.J.Sm.	1	42.6	21.39	47	0	28	0	1	2	1	7	0	9	5
1337	Orchidaceae	*Bulbophyllum comberi* J.J.Verm.	1	42.6	15.75	56	0	26	0	0	0	1	3	0	6	8
1371	Orchidaceae	*Coelogyne dayana* Rchb.f.	1	42.5	17.29	52	8	13	0	0	0	1	5	13	3	4

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1398	Orchidaceae	*Dendrochilum gracilipes* Carr	1	42.2	18.49	51	10	19	0	0	0	0	11	0	4	4
1405	Orchidaceae	*Entomophobia kinabaluensis* Ames de Vogel	1	42.2	20.73	44	5	40	0	0	4	2	0	0	3	2
1334	Orchidaceae	*Bulbophyllum brevicolumna* J.J.Verm.	1	41.8	19.38	50	12	14	0	0	0	0	9	0	7	7
463	Ericaceae	*Rhododendron lanceolatum* Ridl.	1	40.6	22.09	43	13	30	0	0	0	1	5	0	7	2
499	Loganiaceae	*Mitreola sphaerocarpa* Leenh. Leenh.	1	40.6	18.75	51	6	8	0	0	0	4	5	0	10	16
546	Fagaceae	*Castanopsis clemensii* Soepadmo	1	40.5	11.73	67	0	8	0	0	0	0	1	0	7	17
926	Podocarpaceae	*Dacrydium xanthandrum* Pilg.	1	40.5	16.09	52	0	37	0	0	0	0	2	0	7	2
1359	Orchidaceae	*Bulbophyllum tortuosum* Blume Lindl.	1	40.4	20.07	48	0	29	0	0	1	0	12	1	5	4
451	Ericaceae	*Diplycosia salicifolia* Sleumer	1	40.3	20.58	47	18	6	0	0	0	0	4	0	14	10
1322	Orchidaceae	*Appendicula congesta* Ridl.	1	40.2	12.69	63	0	12	0	0	0	0	3	0	11	11
412	Campanulaceae	*Lobelia zeylanica* L.	1	40.1	14	58	0	21	0	0	1	0	6	7	2	6
477	Ericaceae	*Rhododendron suaveolens* Sleumer	1	40.1	10	72	0	11	0	0	0	0	1	0	4	13
1414	Orchidaceae	*Eria robusta Blume* Lindl.	1	40.1	7.54	85	0	6	0	0	0	0	0	0	1	8
1377	Orchidaceae	*Coelogyne planiscapa* Carr	1	39.7	22.11	44	14	19	0	1	0	0	11	0	5	5
1155	Daphniphyllaceae	*Daphniphyllum glaucescens* Blume subsp. *borneense* Stapf T.C.Huang	1	39.3	12.18	60	0	33	0	0	0	0	3	0	1	4
1400	Orchidaceae	*Dendrochilum longipes* J.J.Sm.	1	39	8.29	75	0	23	0	0	0	0	1	0	0	0
1333	Orchidaceae	*Bulbophyllum auratum* Lindl. Rchb.f.	1	38.4	16.73	52	2	22	0	0	1	0	11	0	3	9
1392	Orchidaceae	*Dendrobium kiauense* Ames & C.Schweinf.	1	38.4	23.96	38	14	35	0	0	2	1	0	0	5	3
1342	Orchidaceae	*Bulbophyllum* flavescens Blume Lindl.	1	38.3	21.48	44	7	24	0	0	3	3	7	0	8	5
1363	Orchidaceae	*Chelonistele amplissima* Ames & C.Schweinf. Carr	1	38.3	23.74	42	7	22	0	0	1	4	0	0	8	17
89	Euphorbiaceae	*Aporosa fulvovittata* Schot	1	38.1	11.9	65	0	3	0	0	0	0	3	0	15	14
955	Polygalaceae	*Xanthophyllum tenue* Chodat	1	38	8.58	78	0	0	0	0	0	0	0	0	18	4
868	Araliaceae	*Arthrophyllum borneensis* Philipson Merr.	1	37.9	13.94	56	0	31	0	0	0	0	12	0	0	2
1372	Orchidaceae	*Coelogyne hirtella* J.J.Sm.	1	37.6	10.2	66	1	22	0	0	0	0	1	0	5	5
1270	Moraceae	*Ficus eumorpha* Corner	1	37.4	7.35	81	0	13	0	0	0	0	1	0	0	5
1395	Orchidaceae	*Dendrobium orbiculare* J.J.Sm.	1	37.4	16.51	51	0	24	0	0	0	0	5	0	5	14
1321	Orchidaceae	*Appendicula calcarata* Ridl.	1	37.2	8.22	79	0	3	0	0	0	0	13	0	1	4
414	Campanulaceae	*Pentaphragma aurantiaca* Stapf	1	37	17.22	48	3	37	0	0	0	0	10	0	0	2
1354	Orchidaceae	*Bulbophyllum pocillum* J.J.Verm.	1	36.6	7.05	80	0	19	0	0	0	0	0	0	0	0
583	Fagaceae	*Trigonobalanus verticillata* Forman	1	36.5	24.43	39	6	32	0	0	5	0	0	0	7	11
698	Anacardiaceae	*Rhus borneensis* Stapf	1	36.3	5.19	99	0	0	0	0	0	0	0	0	0	1
1323	Orchidaceae	*Appendicula cornuta* Blume	1	36.3	26.82	36	23	19	0	0	2	0	0	0	8	12
646	Symplocaceae	*Symplocos henschelii* Moric. Benth. ex C.B.Clarke subsp. *henscheli* var. *henscheli*	1	36.1	11.1	60	19	13	0	0	0	0	1	0	7	0
1302	Moraceae	*Ficus setiflora Stapf* var. *adelpha* Corner	1	36	12.81	58	0	23	0	0	0	0	16	0	1	2
511	Staphyleaceae	*Turpinia grandis* B.L.Linden	1	35.9	10.16	69	0	0	0	0	0	0	1	0	16	14
1410	Orchidaceae	*Eria longifolia* Hook.f.	1	35.6	7.79	77	0	11	0	0	0	0	1	1	10	0
481	Ericaceae	*Vaccinium claoxylon* J.J.Sm.	1	35.4	21.08	41	9	31	0	0	4	1	3	0	8	2
449	Ericaceae	*Diplycosia punctulata* Stapf	1	35	17.63	51	0	11	0	0	0	0	28	0	9	1
1404	Orchidaceae	*Dilochia cantleyi* Hook.f. Ridl.	1	34.3	26.49	35	14	33	0	0	7	2	4	0	3	3
1365	Orchidaceae	*Chelonistele sulphurea* Blume Pfitzer var. *sulphurea*	1	34.2	22.69	38	27	13	0	6	0	1	0	0	9	5
962	Sabiaceae	*Meliosma rufopilosa* M.R.Hend.	1	33.9	11.17	63	0	11	0	0	0	0	18	0	5	3
1388	Orchidaceae	*Dendrobium cinereum* J.J.Sm.	1	33.7	23.08	37	16	29	0	0	0	3	0	3	1	10
1415	Orchidaceae	*Eria saccifera* Hook.f.	1	33.5	28.58	34	8	31	0	0	7	1	4	0	6	9
1437	Orchidaceae	*Trichotosia aurea* Ridl. Carr	1	33.4	16.64	46	0	34	0	0	2	2	0	0	11	5
586	Cyperaceae	*Carex cruciata* Wahlenb. var. *cruciata*	1	33.3	17.24	49	0	16	0	0	0	0	8	0	9	19
1368	Orchidaceae	*Coelogyne compressicaulis* Ames & C.Schweinf.	1	33	5.4	90	0	10	0	0	0	0	0	0	0	0
84	Euphorbiaceae	*Aporosa caloneura* Airy Shaw	1	32.2	19.75	40	11	26	0	0	13	0	4	0	4	1
575	Fagaceae	*Lithocarpus papilifer Hatus. ex Soepadmo*	1	32.2	13.2	50	0	39	0	0	3	0	0	0	7	1
1425	Orchidaceae	*Pholidota gibbosa* Blume de Vriese	1	32	27.49	32	15	31	0	0	4	6	1	3	3	6
487	Ericaceae	*Vaccinium phillyreoides* Sleumer	1	31.7	26.02	33	24	25	0	0	7	0	0	0	10	1
454	Ericaceae	*Rhododendron bagobonum* J.J.Copel.	1	31.5	29.87	32	2	28	0	9	7	5	0	0	5	12
1283	Moraceae	*Ficus macilenta* King	1	31.2	6.53	81	1	3	0	0	0	0	4	0	11	0
649	Symplocaceae	*Symplocos laeteviridis* Stapf var. *alternifolia* Noot.	1	30.6	11.88	58	0	7	0	0	0	0	9	0	9	17

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1289	Moraceae	*Ficus oleifolia* King subsp. *intermedia* Corner C.C.Berg	1	30.6	5.54	85	0	13	0	0	0	0	1	0	1	0
920	Magnoliaceae	*Magnolia liliifera* L. Baill. var. *obovata*	1	30.4	27.68	35	9	14	0	0	1	0	14	5	17	5
568	Fagaceae	*Lithocarpus hatusimae* Soepadmo	1	30.2	8.3	67	0	16	0	0	0	0	0	0	4	13
1439	Orchidaceae	*Trichotosia ferox* Blume	1	30.2	28.28	33	1	30	0	0	14	2	7	0	9	5
260	Styracaceae	*Bruinsmia styracoides* Boerl. & Koord.	1	29.8	16.94	45	0	19	0	0	6	5	5	0	10	11
645	Symplocaceae	*Symplocos henschelii* Moric. Benth. ex C.B.Clarke subsp. *henscheli*	1	29.6	17.79	42	23	5	0	0	7	0	0	0	19	3
1349	Orchidaceae	*Bulbophyllum mutabile* Blume Lindl.	1	29.3	4.41	97	0	0	0	0	0	0	3	0	0	0
1012	Rosaceae	*Rubus rosaefolius* Sm.	1	29.2	19.69	42	0	8	0	1	0	4	1	0	8	36
1408	Orchidaceae	*Eria hyacinthoides* Blume Lindl.	1	29.2	13.29	53	0	6	0	0	0	2	2	0	11	25
256	Myricaceae	*Myrica javanica* Blume	1	28.8	25.46	34	4	33	0	0	8	1	8	0	8	3
435	Celastraceae	*Salacia laurifolia* Stapf	1	28.7	19.44	40	0	25	0	0	0	0	18	13	3	1
439	Ericaceae	*Diplycosia barbigera* Sleumer	1	28.7	18.62	38	29	17	0	0	0	0	0	0	10	6
124	Euphorbiaceae	*Homalanthus caloneurus* Airy Shaw	1	28.5	21.22	40	0	17	0	0	1	0	12	0	11	19
652	Symplocaceae	*Symplocos laeteviridis* Stapf var. *laeteviridis*	1	28.5	27.12	34	8	13	0	10	3	1	4	0	22	5
470	Ericaceae	*Rhododendron polyanthemum* Sleumer	1	28.3	33.38	29	14	27	0	0	8	1	12	1	4	5
1197	Myristicaceae	*Horsfieldia montana* Airy Shaw	1	27.9	4.29	94	0	6	0	0	0	0	0	0	0	0
1397	Orchidaceae	*Dendrochilum gibbsiae* Rolfe	1	27.9	17.56	39	5	39	0	0	6	0	0	0	5	6
91	Euphorbiaceae	*Aporosa lagenocarpa* Airy Shaw	1	27.5	8.01	65	0	12	0	0	9	0	0	0	13	0
1401	Orchidaceae	*Dendrochilum muluense* J.J.Wood	1	27.5	3.95	99	0	0	0	0	0	0	0	0	0	1
655	Symplocaceae	*Symplocos laeteviridis* Stapf var. *velutinosa* Noot.	1	27.4	4.33	93	0	6	0	0	0	0	0	0	0	1
1413	Orchidaceae	*Eria nutans* Lindl.	1	27.1	19.19	41	7	5	0	1	0	7	2	1	15	20
442	Ericaceae	*Diplycosia cinnamomifolia* Stapf	1	26.9	4.23	94	0	1	0	0	0	0	0	0	1	3
1339	Orchidaceae	*Bulbophyllum coriaceum* Ridl.	1	26.8	16.79	45	0	12	0	0	0	0	14	0	14	14
82	Euphorbiaceae	*Aporosa basilanensis* Merr.	1	26.7	18.21	42	11	4	0	0	2	0	0	1	19	22
484	Ericaceae	*Vaccinium laurifolium* Blume Miq. var. *sarawakense* Merr. Sleumer	1	26.6	21.22	37	0	24	0	0	3	0	0	12	23	0
1335	Orchidaceae	*Bulbophyllum caudatisepalum* Ames & C.Schweinf.	1	26.4	14.92	43	2	38	0	0	4	0	8	0	4	1
961	Sabiaceae	*Meliosma pinnata* Roxb. Maxim. subsp. *ridleyi* King Beusekom	1	26.3	24.27	33	15	15	0	0	12	2	3	1	17	2
659	Symplocaceae	*Symplocos pendula* Wight var. *hirtistylis* Clarke Noot.	1	25.8	10.26	55	0	17	0	0	4	1	1	1	19	2
907	Menispermaceae	*Stephania reticulata* Forman	1	25.8	16.4	44	0	5	0	0	0	2	2	0	13	33
928	Podocarpaceae	*Phyllocladus hypophyllus* Hook.f.	1	25.8	29.87	29	20	22	0	0	11	1	8	0	6	4
417	Celastraceae	*Celastrus monospermoides* Loes.	1	25.6	18.08	40	0	20	0	0	9	0	1	1	28	1
441	Ericaceae	*Diplycosia ciliolata* Hook.f.	1	25.5	12.2	53	0	6	0	0	0	0	18	0	13	10
911	Araucariaceae	*Agathis orbicula* de Laub.	1	25.5	21.04	37	13	4	0	2	0	7	1	0	15	20
491	Juglandaceae	*Engelhardtia rigida* Blume	1	25.4	24.37	36	1	1	0	0	0	2	3	2	21	34
1360	Orchidaceae	*Bulbophyllum unguiculatum* Rchb.f.	1	25.3	7.36	65	0	19	0	0	1	0	0	0	15	1
639	Symplocaceae	*Symplocos anomala* Brand	1	25.2	5.81	78	0	0	0	0	0	0	0	0	22	0
515	Clethraceae	*Clethra pachyphylla* Merr.	1	25.1	5.65	77	0	4	0	0	0	0	1	3	15	0
581	Fagaceae	*Quercus lowii* King	1	25	15.3	42	0	27	0	0	6	0	0	0	21	3
1423	Orchidaceae	*Pholidota carnea* Blume Lindl. var. *carnea*	1	25	14.97	42	0	28	0	5	3	1	5	6	9	0
1438	Orchidaceae	*Trichotosia brevipedunculata* Ames & C.Schweinf.	1	25	11.45	54	0	1	0	0	0	0	17	0	14	13
647	Symplocaceae	*Symplocos henschelii* Moric. Benth. ex C.B.Clarke var. *maingayi* Benth. ex C.B.Clarke Noot.	1	24.7	9.37	58	3	0	0	0	0	4	0	0	20	14
882	Liliaceae	*Ophiopogon caulescens* Blume Backer	1	24.5	4.3	89	0	1	0	0	0	0	0	0	3	7
623	Hypericaceae	*Hypericum japonicum* Thunb. ex A.Murray	1	24.2	20.94	37	0	12	0	1	1	1	8	0	12	28
572	Fagaceae	*Lithocarpus luteus* Soepadmo	1	24.1	6.77	70	0	0	0	0	0	0	0	0	30	0
382	Flacourtiaceae	*Xylosma sumatrana* Sloot.	1	24	11.62	51	0	5	0	0	0	0	3	0	14	26
447	Ericaceae	*Diplycosia memecyloides* Stapf	1	23.9	11.48	52	0	6	0	0	0	0	20	0	12	9
560	Fagaceae	*Lithocarpus confertus* Soepadmo	1	23.7	9.26	54	0	27	0	0	4	0	0	0	15	0
459	Ericaceae	*Rhododendron durionifolium* Becc.	1	23.6	27.87	27	26	19	0	3	16	1	0	0	6	2
243	Flagellariaceae	*Joinvillea ascendens* Gaudich. ex Brongn. & Gris subsp. *borneensis* Becc. Newell	1	23.1	10.6	49	0	24	0	0	7	1	4	1	7	6
1418	Orchidaceae	*Geesinkorchis alaticallosa* de Vogel	1	23.1	14	46	0	1	0	2	0	1	11	0	13	25

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1318	Orchidaceae	*Agrostophyllum glumaceum* Hook.f.	1	23	26.23	32	16	1	0	0	1	1	0	0	30	19
318	Burseraceae	*Santiria grandiflora* Kalkman	1	22	18.76	37	12	3	0	0	0	5	1	0	20	22
1089	Meliaceae	*Dysoxylum macrocarpum* Blume	1	22	22.65	32	23	4	0	9	0	2	0	2	17	11
651	Symplocaceae	*Symplocos laeteviridis* Stapf var. *kinabaluensis* Heine Noot.	1	21.8	4.23	85	0	1	0	0	0	0	11	0	2	0
418	Celastraceae	*Euonymus castaneifolius* Ridl.	1	21.7	38.04	24	22	7	0	0	6	1	6	10	16	7
469	Ericaceae	*Rhododendron pneumonanthum* Sleumer	1	21.7	3.29	95	0	4	0	0	0	0	0	0	0	0
1374	Orchidaceae	*Coelogyne kinabaluensis* Ames & C.Schweinf.	1	21.7	3.37	94	0	5	0	0	0	0	1	0	0	1
949	Polygalaceae	*Xanthophyllum purpureum* Ridl.	1	21.5	29.94	29	9	2	1	0	1	3	11	2	28	13
1369	Orchidaceae	*Coelogyne craticulaelabris* Carr	1	21.5	22.75	33	0	20	0	0	1	0	22	11	9	6
1061	Meliaceae	*Aglaia monozyga* Harms	1	21.3	15.55	42	0	3	0	0	0	4	2	0	36	13
235	Dilleniaceae	*Saurauia tristyla* DC.	1	21.2	10.85	50	0	2	0	11	0	0	9	0	15	13
1005	Rosaceae	*Prunus spicata* Kalkman	1	21.2	19.77	35	0	22	0	0	1	0	20	2	16	4
1183	Loranthaceae	*Taxillus chinensis* DC. Danser	1	21	13.02	45	0	6	0	0	0	0	11	4	14	20
1327	Orchidaceae	*Appendicula torta* Blume	1	20.7	18.4	35	2	13	0	9	14	0	7	0	13	7
1370	Orchidaceae	*Coelogyne cuprea* H.Wendl. & Kraenzl.	1	20.7	5.7	71	0	1	0	0	0	0	12	0	12	5
637	Symplocaceae	*Symplocos adenophylla* Wall. & G.Don	1	20.4	8.88	53	0	9	1	1	5	0	8	4	19	0
551	Fagaceae	*Castanopsis javanica* Blume A.DC.	1	20.3	15.75	40	1	1	0	0	0	0	0	13	20	25
476	Ericaceae	*Rhododendron stenophyllum* Hook.f.	1	20.2	5.23	72	0	2	0	0	0	0	0	0	24	1
450	Ericaceae	*Diplycosia rufa* Stapf	1	20.1	29.63	28	0	12	0	11	1	0	15	1	20	13
1390	Orchidaceae	*Dendrobium cymbulipes* J.J.Sm.	1	20	2.86	100	0	0	0	0	0	0	0	0	0	0
885	Liliaceae	*Pleomele angustifolia* Roxb. N.E.Br.	1	19.8	16.24	39	0	5	0	0	1	0	1	11	35	8
710	Labiatae	*Paraphlomis javanica* Blume Prain	1	19.7	12.87	45	0	1	0	1	0	1	9	0	11	31
276	Thymelaeaceae	*Wikstroemia androesaenifolia* Decne.	1	19.4	21.67	32	0	14	0	0	9	1	8	0	23	13
1373	Orchidaceae	*Coelogyne incrassata* Blume Lindl.	1	19.4	14.61	39	0	24	1	0	6	0	13	0	14	2
869	Araliaceae	*Arthrophyllum collinum* Philipson	1	19.2	10.47	48	0	10	0	0	0	0	1	0	30	11
711	Labiatae	*Plectranthus galeatus* Vahl	1	19	4.45	76	0	6	0	0	0	0	0	0	18	0
930	Podocarpaceae	*Podocarpus laubenfelsii* Tiong	1	19	15.53	39	0	4	0	2	0	1	6	0	12	36
648	Symplocaceae	*Symplocos johniana* Stapf	1	18.8	2.86	96	0	0	0	0	0	0	0	0	4	0
544	Violaceae	*Viola sumatrana* Miq.	1	18.6	10.2	49	0	3	0	0	0	0	21	0	14	12
1205	Myristicaceae	*Horsfieldia xanthina* Airy Shaw subsp. *xanthina*	1	18.6	18.62	32	26	5	0	0	0	2	2	1	19	14
1291	Moraceae	*Ficus parietalis* Blume	1	18.6	31.53	26	4	1	2	2	3	3	6	13	25	16
1340	Orchidaceae	*Bulbophyllum deltoideum* Ames & C.Schweinf.	1	18.4	4.99	65	0	28	0	0	1	0	4	0	0	1
161	Euphorbiaceae	*Mallotus lancifolius* Hook.f.	1	18.3	12.36	44	0	0	0	0	0	0	15	0	19	21
1238	Nepenthaceae	*Nepenthes tentaculata* Hook.f.	1	18.1	22.5	28	1	24	2	2	20	3	2	1	8	9
1406	Orchidaceae	*Eria farinosa* Ames & C.Schweinf.	1	18.1	3.07	91	0	1	0	0	0	0	6	0	0	2
1402	Orchidaceae	*Dendrochilum oxylobum* Schltr.	1	18	17.38	35	11	1	0	0	0	4	3	2	15	30
1301	Moraceae	*Ficus setiflora* Stapf	1	17.8	3.84	79	0	7	0	0	0	0	14	0	0	0
281	Umbelliferae	*Hydrocotyle javanica* Thunb.	1	17.7	24.41	27	0	25	0	0	9	0	5	12	18	3
1325	Orchidaceae	*Appendicula foliosa* Ames & Schltr.	1	17.6	12.39	42	0	1	0	0	0	0	9	15	26	7
1431	Orchidaceae	*Spathoglottis microchilina* Kraenzl.	1	17.6	10.9	46	0	0	0	0	0	0	1	0	30	23
1177	Loranthaceae	*Macrosolen flammeus* Danser	1	17.5	21.24	31	0	7	5	7	0	0	12	9	21	8
112	Euphorbiaceae	*Bridelia glauca* Blume var. *sosopodonica* Airy Shaw S.Dressler	1	17.3	23.41	27	0	8	27	0	3	2	1	0	12	18
1421	Orchidaceae	*Oberonia affinis* Ames & C.Schweinf.	1	17.3	10.03	48	0	3	0	0	0	0	19	0	15	15
1180	Loranthaceae	*Macrosolen tubiflorus* Ridl. Danser	1	17.1	11.18	44	0	11	0	0	3	0	9	0	31	3
1430	Orchidaceae	*Podochilus tenuis* Lindl.	1	17.1	6.24	61	0	0	0	0	0	0	0	0	32	8
195	Caprifoliaceae	*Viburnum hispidulum* Kern	1	16.7	11.79	43	0	1	0	0	0	0	9	3	27	17
588	Cyperaceae	*Carex perakensis* C.B.Clarke var. *borneensis* C.B.Clarke Noot.	1	16.7	19	30	0	21	0	0	11	5	6	13	13	2
128	Euphorbiaceae	*Macaranga angulata* S.J.Davies	1	16.6	2.36	100	0	0	0	0	0	0	0	0	0	0
873	Araliaceae	*Schefflera avensis* Harms	1	16.6	26.63	26	0	6	0	0	14	8	2	19	19	5
1352	Orchidaceae	*Bulbophyllum obtusum* Blume Lindl.	1	16.6	2.36	100	0	0	0	0	0	0	0	0	0	0
10	Annonaceae	*Goniothalamus clemensii* Bân	1	16.5	12.94	40	0	7	0	0	1	0	16	0	12	24
1070	Meliaceae	*Aglaia squamulosa* King	1	16.4	34.13	23	0	21	0	0	20	1	8	0	22	5
559	Fagaceae	*Lithocarpus clementiana* King ex Hook.f. A.Camus	1	16.3	8.58	48	0	9	0	0	0	0	1	12	30	0
1383	Orchidaceae	*Coelogyne venusta* Rolfe	1	16.3	2.6	92	0	8	0	0	0	0	0	0	0	0
1427	Orchidaceae	*Pholidota pectinata* Ames	1	16.3	2.33	99	0	0	0	0	0	0	0	0	0	1

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
468	Ericaceae	*Rhododendron orbiculatum* Ridl.	1	16.2	9.92	39	34	15	0	0	7	0	0	0	5	1
924	Podocarpaceae	*Dacrycarpus* imbricatus Blume de Laub. var. *patulus* de Laub.	1	16.2	22.91	28	0	17	0	19	4	0	8	4	16	4
1316	Moraceae	*Streblus glaber* Merr. Corner	1	16.2	24.71	26	0	22	0	0	12	0	4	14	10	13
1317	Orchidaceae	*Adenoncos parviflora* Ridl.	1	16.1	10.07	46	0	0	0	0	0	0	3	0	19	32
1331	Orchidaceae	*Bulbophyllum acutum* J.J.Sm.	1	16.1	2.31	99	0	0	0	0	0	0	0	0	0	1
194	Caprifoliaceae	*Viburnum clemensae Kern*	1	16	2.38	97	0	1	0	0	0	0	0	0	2	0
644	Symplocaceae	*Symplocos fasciculata* Zoll.	1	15.9	57.83	17	14	5	0	2	5	10	6	12	16	13
657	Symplocaceae	*Symplocos ophirensis* C.B.Clarke subsp. *cumingiana* Brand Noot.	1	15.8	5.35	63	0	1	0	0	0	0	0	0	36	0
579	Fagaceae	*Quercus argentata* Korth.	1	15.2	13.36	38	0	6	0	2	3	0	17	1	23	11
570	Fagaceae	*Lithocarpus lampadarius* Gamble A.Camus	1	15	4.73	64	0	6	0	0	3	0	0	0	26	1
191	Burmaniaceae	*Burmannia longifolia* Becc.	1	14.8	17.38	30	0	23	0	0	20	3	14	0	5	4
1382	Orchidaceae	*Coelogyne subintegra* J.J.Sm.	1	14.8	15.66	35	0	0	0	0	0	0	16	10	28	11
910	Araucariaceae	*Agathis lenticula* de Laub.	1	14.7	11.29	41	0	2	0	1	0	0	2	0	14	40
1243	Moraceae	*Artocarpus excelsus* F.M.Jarrett	1	14.7	13.12	38	0	4	0	1	0	0	7	0	14	36
444	Ericaceae	*Diplycosia commutata* Sleumer	1	14.4	2.09	99	0	0	0	0	0	0	1	0	0	1
460	Ericaceae	*Rhododendron fallacinum* Sleumer	1	14.3	2.94	82	0	0	0	0	0	0	0	0	6	12
811	Dipterocarpaceae	*Shorea monticola* P.S.Ashton	1	14.1	11.55	35	7	21	0	0	31	0	0	0	7	0
1387	Orchidaceae	*Dendrobium bifarium* Lindl. ex Hook.f.	1	14	5.14	60	0	0	0	0	0	0	0	0	4	36
149	Euphorbiaceae	*Macaranga rostrata* Heine	1	13.9	6.7	51	0	4	0	0	3	2	1	0	17	21
446	Ericaceae	*Diplycosia heterophylla* Blume var. *latifolia* Blume Sleumer	1	13.9	14.78	31	17	5	0	24	8	0	1	3	12	0
516	Haloragaceae	*Haloragis micrantha* Thunb. R.Br. ex Siebold & Zucc.	1	13.9	12.91	37	0	2	0	0	0	0	12	7	24	18
146	Euphorbiaceae	*Macaranga petanostyla* Airy Shaw	1	13.8	2.94	81	0	0	0	0	0	0	10	0	1	8
480	Ericaceae	*Vaccinium bancanum* Miq. var. tenuinervium J.J.Sm.	1	13.8	19.55	26	20	8	0	3	19	0	1	2	16	5
587	Cyperaceae	*Carex filicina* Nees	1	13.6	14.25	35	0	3	0	0	0	0	27	4	24	7
1329	Orchidaceae	*Bromheadia cecieliae* Kruizinga	1	13.6	2.15	93	0	6	0	0	0	0	1	0	0	1
921	Magnoliaceae	*Magnolia macklottii* Korth. Dandy *var. macklottii*	1	13.4	2.02	95	0	5	0	0	0	0	0	0	0	0
547	Fagaceae	*Castanopsis costata* Blume A.DC.	1	13.2	8.88	43	9	0	0	6	0	0	0	0	35	7
1351	Orchidaceae	*Bulbophyllum nematocaulon* Ridl.	1	13.2	2.24	88	0	11	0	0	0	0	0	0	0	1
1426	Orchidaceae	*Pholidota imbricata* Hook.	1	12.8	11.44	36	0	27	0	0	3	1	24	0	8	0
827	Dipterocarpaceae	*Shorea platyclados* Slooten ex Endert	1	12.7	16.98	28	9	10	0	1	4	7	0	1	19	21
277	Thymelaeaceae	*Wikstroemia brachyantha* Merr.	1	12.6	2.56	82	0	0	0	0	0	0	6	0	1	11
925	Podocarpaceae	*Dacrycarpus kinabaluensis* Wasscher de Laub.	1	12.5	3.23	73	0	0	0	0	0	0	8	0	9	10
929	Podocarpaceae	*Podocarpus brevifolius* Stapf Foxw.	1	12.5	3.23	73	0	0	0	0	0	0	8	0	9	10
1348	Orchidaceae	*Bulbophyllum montense* Ridl.	1	12.5	5.91	53	0	0	0	0	0	0	0	0	5	43
1378	Orchidaceae	*Coelogyne planiscapa* Carr var. *planiscapa*	1	12.5	1.77	100	0	0	0	0	0	0	0	0	0	0
888	Liliaceae	*Smilax lanceifolia* Roxb.	1	12.4	1.83	98	0	0	0	0	0	0	0	0	2	0
1268	Moraceae	*Ficus disticha* Blume subsp. *disticha*	1	12.2	19.26	26	1	11	0	23	15	0	8	3	13	0
889	Liliaceae	*Smilax leucophylla* Blume	1	11.9	24.12	23	4	12	0	0	20	2	0	4	22	13
965	Sabiaceae	*Sabia racemosa* Chen subsp. *kinabaluensis* Water	1	11.8	16.43	30	0	2	0	0	0	1	12	10	16	30
1003	Rosaceae	*Prunus laxinervis* Kalkman	1	11.8	18.83	28	0	2	0	0	0	0	21	14	20	15
431	Celastraceae	*Microtropis platyphylla* Merr.	1	11.6	10.77	37	0	0	0	0	0	0	0	2	29	32
510	Staphyleaceae	*Turpinia borneensis* Merr. & L.M.Perry B.L.Linden	1	11.6	1.67	99	0	0	0	0	0	0	0	0	0	1
1422	Orchidaceae	*Oberonia kinabaluensis* Ames & C.Schweinf.	1	11.5	1.63	100	0	0	0	0	0	0	0	0	0	0
462	Ericaceae	*Rhododendron jasminiflorum* Hook. var. *oblongifolium* Sleumer	1	11.4	8.39	39	0	1	0	0	0	20	22	18	0	0
825	Dipterocarpaceae	*Shorea parvistipulata* Heim subsp. *nebulosa* Meijer Ashton	1	11.4	7.74	44	0	0	0	0	0	0	7	2	18	28
471	Ericaceae	*Rhododendron praetervisum* Sleumer	1	11.1	1.73	95	0	0	0	0	0	0	0	0	0	5
874	Araliaceae	*Schefflera bipalmatifolia* Merr.	1	11.1	1.57	100	0	0	0	0	0	0	0	0	0	0
974	Mimosaceae	*Acacia donnaiensis* Gagnep.	1	11.1	9.4	39	0	0	0	0	0	0	2	0	23	36
512	Staphyleaceae	*Turpinia nitida* Merr. & L.M.Perry	1	10.9	13.37	32	0	1	0	0	0	0	7	17	21	22
448	Ericaceae	*Diplycosia pinifolia* Stapf	1	10.7	11.64	35	0	1	0	0	0	0	23	0	18	23
902	Menispermaceae	*Cyclea elegans* King	1	10.6	4.27	58	0	0	0	0	0	0	1	0	21	20

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1282	Moraceae	*Ficus leptocalama* Corner	1	10.6	4.24	57	0	0	0	0	0	0	0	0	9	33
1407	Orchidaceae	*Eria floribunda* Lindl.	1	10.6	3.57	64	0	0	0	0	0	0	11	0	2	23
1306	Moraceae	*Ficus tarennifolia* Corner	1	10.5	12.63	33	0	1	0	0	0	1	25	0	19	21
1435	Orchidaceae	*Trichoglottis tinekeae* Schuit.	1	10.5	1.48	100	0	0	0	0	0	0	0	0	0	0
580	Fagaceae	*Quercus lineata* Blume	1	10.3	11.34	32	0	15	0	9	1	0	7	11	19	7
445	Ericaceae	*Diplycosia fimbriata* Sleumer	1	10.2	1.68	93	0	0	0	0	0	0	4	0	3	0
1332	Orchidaceae	*Bulbophyllum aeolium* Ames	1	10.1	1.43	100	0	0	0	0	0	0	0	0	0	0
899	Chloranthaceae	*Ascarina philippinensis* C.B.Rob.	1	9.8	1.5	96	0	0	0	0	0	0	0	0	4	0
192	Caprifoliaceae	*Sambucus javanica* Reinw. ex Blume	1	9.4	1.6	90	0	0	0	0	0	0	0	0	0	10
223	Dilleniaceae	*Saurauia actinidiifolia* Stapf	1	9.3	6.27	44	0	0	0	0	0	0	0	0	23	33
653	Symplocaceae	*Symplocos laeteviridis* Stapf var. *mjobergii* Merr. Noot.	1	9.3	1.36	98	0	2	0	0	0	0	0	0	0	0
1343	Orchidaceae	*Bulbophyllum fulvibulbum* J.J.Verm.	1	9.2	1.33	99	0	0	0	0	0	0	0	0	0	1
1380	Orchidaceae	*Coelogyne radioferens* Ames & C.Schweinf.	1	9.2	1.63	89	0	0	0	0	0	0	0	0	5	6
282	Umbelliferae	*Sanicula elata* Buch.-Ham. ex D.Don	1	9	1.29	99	0	0	0	0	0	0	0	0	0	1
242	Flagellariaceae	*Hanguana major* Airy Shaw	1	8.9	7.51	40	0	0	0	0	0	0	20	0	18	21
607	Cyperaceae	*Scirpus subcapitatus* Thwaites	1	8.9	4.71	51	0	0	0	0	0	0	12	0	26	12
964	Sabiaceae	*Sabia parviflora* Wall. subsp. *parviflora*	1	8.8	7.5	40	0	0	0	0	0	0	21	0	19	21
465	Ericaceae	*Rhododendron maxwellii* Gibbs	1	8.7	1.45	92	0	0	0	0	0	0	1	0	0	8
457	Ericaceae	*Rhododendron buxoides* Sleumer	1	8.5	1.84	79	0	0	0	0	0	0	0	0	1	19
1174	Loranthaceae	*Lepidaria sabaensis* Stapf Tiegh.	1	8.5	12.99	29	0	0	0	0	0	0	15	9	25	22
1412	Orchidaceae	*Eria major* Ridl. ex Stapf	1	8.5	6.95	40	0	0	0	0	0	0	19	0	20	20
392	Proteaceae	*Helicia pterygota* Sleumer	1	8.4	4.09	53	0	0	0	0	0	0	1	0	35	11
867	Araliaceae	*Aralia merrillii* C.B.Shang	1	8.4	7.6	39	0	0	0	0	0	0	21	0	18	22
443	Ericaceae	*Diplycosia clementium* Sleumer	1	8.3	1.4	91	0	0	0	0	0	0	1	0	0	8
452	Ericaceae	*Diplycosia urceolata* Stapf	1	8.3	10.45	33	0	0	0	2	0	0	29	2	25	10
997	Rosaceae	*Photinia davidiana* Decne. Cardot	1	8.3	1.4	91	0	0	0	0	0	0	1	0	0	8
1411	Orchidaceae	*Eria magnicallosa* Ames & C.Schweinf.	1	8.3	1.49	88	0	0	0	0	0	0	1	0	0	11
453	Ericaceae	*Gaultheria borneensis* Stapf	1	8.2	1.31	94	0	0	0	0	0	0	1	0	0	5
475	Ericaceae	*Rhododendron rugosum* Low ex Hook.f.	1	8.2	1.17	100	0	0	0	0	0	0	0	0	0	0
483	Ericaceae	*Vaccinium cordifolium* Stapf	1	8.2	1.31	94	0	0	0	0	0	0	1	0	0	5
933	Podocarpaceae	*Prumnopitys amarus* Blume de Laub.	1	8.2	1.31	94	0	0	0	0	0	0	1	0	0	5
283	Umbelliferae	*Trachymene saniculaefolia* Stapf	1	7.9	1.12	100	0	0	0	0	0	0	0	0	0	0
411	Campanulaceae	*Lobelia borneensis* Hemsl. Moeliono	1	7.9	1.43	88	0	0	0	0	0	0	1	0	0	12
474	Ericaceae	*Rhododendron retivenium* Sleumer	1	7.9	1.12	100	0	0	0	0	0	0	0	0	0	0
611	Balanophoraceae	*Balanophora lowii* Hook.f.	1	7.9	1.2	97	0	0	0	0	0	0	0	0	0	3
1355	Orchidaceae	*Bulbophyllum puntjakense* J.J.Sm.	1	7.9	1.12	100	0	0	0	0	0	0	0	0	0	0
251	Gnetaceae	*Gnetum latifolium* Blume var. *minus* Foxw. Markgr.	1	7.8	4.93	45	0	0	0	0	0	0	0	0	14	40
440	Ericaceae	*Diplycosia chrysothrix* Stapf	1	7.8	7.1	38	0	0	0	0	0	0	0	0	29	33
640	Symplocaceae	*Symplocos buxifolia* Stapf	1	7.7	1.1	100	0	0	0	0	0	0	0	0	0	0
1434	Orchidaceae	*Thrixspermum triangulare* Ames & C.Schweinf.	1	7.5	3.03	57	0	0	0	0	0	0	0	0	9	34
1366	Orchidaceae	*Cleisocentron merrillianum* Ames Christenson	1	7.1	2.73	58	0	0	0	0	0	0	0	0	0	42
1436	Orchidaceae	*Trichotosia aporina* Hook.f. Kraenzl.	1	7	1.95	69	0	0	0	0	0	0	0	0	1	30
472	Ericaceae	*Rhododendron quadrasianum* S.Vidal var. *cuneifolium* Stapf Cop. f.	1	6.9	1.28	86	0	0	0	0	0	0	0	0	0	14
1336	Orchidaceae	*Bulbophyllum ceratostylis* J.J.Sm.	1	6.9	0.98	100	0	0	0	0	0	0	0	0	0	0
627	Bignoniaceae	*Radermachera ramiflora* Steenis	1	6.8	12.02	27	0	0	0	0	0	2	24	2	18	27
1379	Orchidaceae	*Coelogyne plicatissima* Ames & C.Schweinf.	1	6.6	2.68	57	0	0	0	0	0	0	0	0	0	43
569	Fagaceae	*Lithocarpus havilandii* Stapf Barnett	1	6.5	0.92	100	0	0	0	0	0	0	0	0	0	0
960	Sabiaceae	*Meliosma lepidota* Blume subsp. *kinabaluensis* Beusekom	1	6.5	3.28	52	0	0	0	0	0	0	4	0	32	13
901	Elaeagnaceae	*Elaeagnus triflora* Roxb.	1	6.3	1.57	74	0	0	0	0	0	0	0	0	26	0
304	Burseraceae	*Canarium pilosum* Benn. subsp. *borneensis* Leenh.	1	6.2	3.92	46	0	0	0	0	0	0	12	0	34	8
916	Magnoliaceae	*Magnolia carsonii* Dandy ex Noot. var. *carsonii*	1	6.2	0.87	100	0	0	0	0	0	0	0	0	0	0
1350	Orchidaceae	*Bulbophyllum mutabile* Blume Lindl. var. *obesum* J.J.Verm.	1	6	0.92	95	0	0	0	0	0	0	0	0	0	5
1152	Boraginaceae	*Cordia dichotoma* G.Forst.	1	5.8	2.51	55	0	0	0	12	0	0	0	0	33	0

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1367	Orchidaceae	*Coelogyne clemensii* Ames & C.Schweinf.	1	5.8	1.14	83	0	0	0	0	0	0	0	0	0	17
1269	Moraceae	*Ficus endospermifolia* Corner	1	5.7	0.8	100	0	0	0	0	0	0	0	0	0	0
918	Magnoliaceae	*Magnolia lasia* Noot.	1	5.4	1.29	75	0	0	0	0	0	0	0	0	4	21
1346	Orchidaceae	*Bulbophyllum microglossum* Ridl.	1	5.3	0.75	100	0	0	0	0	0	0	0	0	0	0
1326	Orchidaceae	*Appendicula minutiflora* Ames & Schweinf.	1	4.9	0.72	98	0	0	0	0	0	0	0	0	0	2
280	Umbelliferae	*Eryngium foetidum* L.	1	3.9	0.55	100	0	0	0	0	0	0	0	0	0	0
1164	Loranthaceae	*Helixanthera apodanthes* Danser	1	3.8	1.56	57	0	0	0	0	0	0	0	0	26	17
1393	Orchidaceae	*Dendrobium kurashigei* Yukawa	1	2.8	0.4	100	0	0	0	0	0	0	0	0	0	0
922	Magnoliaceae	*Magnolia persuaveolens* Dandy subsp. *rigida* Noot.	1	2.6	0.37	100	0	0	0	0	0	0	0	0	0	0
1362	Orchidaceae	*Calanthe sylvatica* Thouars Lindl.	1	2.5	0.35	100	0	0	0	0	0	0	0	0	0	0
96	Euphorbiaceae	*Aporosa rhacostyla* Airy Shaw	2	61.5	9.23	6	63	16	0	0	7	1	1	0	6	0
792	Dipterocarpaceae	*Shorea flaviflora* G.H.S.Wood ex P.S.Ashton	2	56.7	12.3	14	58	6	0	2	7	0	1	0	10	1
833	Dipterocarpaceae	*Shorea sagittata* Ashton	2	56.1	12.21	5	56	12	0	3	5	1	3	0	13	1
45	Annonaceae	*Polyalthia hookeriana* King	2	53	11.22	3	59	2	0	2	2	0	0	5	22	4
807	Dipterocarpaceae	*Shorea macroptera* Dyer subsp. *bailloni*	2	52.6	11.48	0	55	7	0	7	14	1	0	3	14	0
1223	Myristicaceae	*Knema riangensis* W.J.de Wilde	2	50	15.13	11	50	14	0	0	4	0	12	0	9	0
878	Araliaceae	*Schefflera multiarcuata* Frodin	2	48	14.22	7	48	20	0	0	9	2	0	1	13	0
14	Annonaceae	*Goniothalamus uvarioides* King	2	47.9	10.23	1	57	9	0	0	7	2	4	1	19	1
1260	Moraceae	*Ficus bruneiensis* Corner	2	46.5	14.82	8	46	22	0	0	8	1	1	0	12	0
680	Anacardiaceae	*Gluta laxiflora* Ridl.	2	45.9	15.19	6	49	7	0	2	2	3	5	0	22	3
848	Dipterocarpaceae	*Vatica coriacea* P.S.Ashton	2	45.6	15.03	2	46	6	18	0	5	1	1	0	18	2
369	Flacourtiaceae	*Hydnocarpus sumatrana* Miq. Koord. var. *pentagyna* Slooten Sleumer	2	44.2	18.62	8	44	7	0	1	9	1	5	1	23	1
729	Dipterocarpaceae	*Dipterocarpus nudus* Vesque	2	44	7.85	0	61	0	0	21	1	0	0	3	14	0
800	Dipterocarpaceae	*Shorea iliasii* P.S.Ashton	2	43.8	15.48	0	46	3	0	9	12	6	2	3	19	0
600	Cyperaceae	*Mapania cuspidata* Miq. Uittien var. *cuspidata*	2	43.2	17.4	0	46	2	0	0	11	1	14	0	24	1
374	Flacourtiaceae	*Ryparosa glauca* Ridl.	2	42.6	16.21	7	43	21	0	0	14	2	1	3	10	0
98	Euphorbiaceae	*Aporosa stenostachys* Airy Shaw	2	42.1	16.31	0	44	3	0	8	8	6	2	3	21	3
127	Euphorbiaceae	*Macaranga aetheadenia* Airy Shaw	2	41.6	14.88	1	47	4	0	0	18	0	10	0	16	4
743	Dipterocarpaceae	*Dryobalanops oblongifolia* Dyer	2	41.5	15.75	5	41	24	0	1	12	7	1	2	7	0
455	Ericaceae	*Rhododendron brookeanum* Lowe & Lindl. var. *brookeanum*	2	40.5	20.42	23	42	3	0	0	9	0	0	0	22	1
1011	Rosaceae	*Rubus moluccanus* L. var. *discolor* Blume Kalkman	2	39.9	19.04	19	43	2	0	0	7	0	1	2	24	1
1396	Orchidaceae	*Dendrobium secundum* Blume Lindl.	2	39.5	13.86	12	48	8	0	0	0	0	4	6	20	1
272	Thymelaeaceae	*Gonystylus nervosus* Airy Shaw	2	39.4	20.81	13	41	4	0	6	2	0	5	0	25	4
813	Dipterocarpaceae	*Shorea myrionerva* Symington	2	39.3	19.42	13	39	20	0	0	7	2	0	1	13	5
832	Dipterocarpaceae	*Shorea rubra* P.S.Ashton	2	38.4	11.75	2	50	0	2	0	16	2	4	2	21	0
1298	Moraceae	*Ficus rubromidotis* Corner	2	37.9	21.13	7	38	13	0	0	14	1	12	0	14	0
704	Anacardiaceae	*Swintonia foxworthyi* Elmer	2	37.4	17.01	13	44	3	0	2	1	2	3	0	29	3
151	Euphorbiaceae	*Macaranga trachyphylla* Airy Shaw	2	37	14.81	0	44	1	0	0	10	10	1	7	26	0
746	Dipterocarpaceae	*Hopea bracteata* Burck	2	36.9	22.62	4	37	9	0	2	5	4	8	1	25	5
1215	Myristicaceae	*Knema luteola* W.J.de Wilde	2	36.4	13.79	9	42	27	0	0	15	0	4	0	2	0
467	Ericaceae	*Rhododendron nieuwenhuisii* J.J.Sm.	2	36.3	19.25	11	37	27	0	0	13	2	0	0	10	0
722	Dipterocarpaceae	*Dipterocarpus elongatus* Korth.	2	36.3	16.23	2	43	1	0	5	3	5	0	4	27	10
1000	Rosaceae	*Prunus arborea* Blume Kalkman var. *stipulacea* King Kalkman	2	36.3	24.44	21	38	1	0	2	1	0	1	0	31	6
1225	Myristicaceae	*Knema subhirtella* W.J.de Wilde	2	35.7	23.19	7	37	5	0	5	4	3	1	3	28	9
703	Anacardiaceae	*Swintonia acuta* Engl.	2	35.5	23.7	3	36	2	9	0	6	1	1	2	28	13
791	Dipterocarpaceae	*Shorea ferruginea* Dyer ex Brandis	2	35.2	25.62	9	35	2	0	5	11	1	4	1	24	7
764	Dipterocarpaceae	*Hopea treubii* Heim	2	35.1	18.48	0	39	1	0	13	9	5	2	7	24	0
153	Euphorbiaceae	*Macaranga umbrosa* S.J.Davies	2	34.9	23.52	13	35	14	0	9	8	0	2	2	16	1
489	Ericaceae	*Vaccinium uniflorum* J.J.Sm.	2	34.9	16.16	24	44	6	0	0	0	0	3	0	21	2
1213	Myristicaceae	*Knema linguiformis* Sinclair W.J.de Wilde	2	34.3	23.32	16	36	1	2	11	1	0	0	3	28	1
1075	Meliaceae	*Chisocheton macranthus* Merr. Airy Shaw	2	34.1	12.92	5	48	0	0	0	6	0	4	1	35	2
394	Rhizophoraceae	*Anisophyllea beccariana* Baill.	2	34	24.08	6	34	3	1	10	11	3	1	9	20	3
789	Dipterocarpaceae	*Shorea falcifera* Dyer	2	34	22.09	7	34	14	2	8	8	2	1	8	15	0

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
860	Dipterocarpaceae	*Vatica odorata* Griff. Symington subsp. *mindanensis* Foxw. P.S.Ashton	2	33.7	26.49	16	34	5	0	8	10	0	0	1	19	8
782	Dipterocarpaceae	*Shorea crassa* P.S.Ashton	2	33.2	18.57	0	38	4	5	0	15	3	9	0	24	1
226	Dilleniaceae	*Saurauia amoena* Stapf	2	33	14.96	22	44	5	0	0	8	0	0	3	16	3
872	Araliaceae	*Schefflera avenis* Miq. Harms	2	33	24.71	18	33	10	0	21	4	0	0	1	12	0
1284	Moraceae	*Ficus macrostyla* Corner	2	33	17.19	0	40	0	0	0	19	4	14	4	18	1
364	Flacourtiaceae	*Hydnocarpus calophyllus* Ridl. Sleumer	2	32.6	25.21	6	33	15	0	0	19	1	4	0	19	4
44	Annonaceae	*Polyalthia flagellaris* Becc. Airy Shaw	2	32.3	25.87	5	33	5	0	0	25	3	3	0	22	4
788	Dipterocarpaceae	*Shorea faguetioides* P.S.Ashton	2	32	25.45	5	32	13	0	0	16	6	2	1	22	3
423	Celastraceae	*Lophopetalum glabrum* Ding Hou	2	31.9	24.44	11	34	4	0	0	14	1	4	0	24	7
79	Euphorbiaceae	*Antidesma riparium* Airy Shaw	2	31.5	14.17	15	45	1	0	0	1	0	0	3	28	7
1319	Orchidaceae	*Agrostophyllum trifidum* Schltr.	2	31.4	20.02	20	36	12	0	0	4	7	2	2	9	9
556	Fagaceae	*Lithocarpus bennettii* Miq. Rehder	2	31.3	22.28	21	37	5	0	0	1	0	6	1	30	0
352	Flacourtiaceae	*Casearia hosei* Merr.	2	31.1	16.08	21	41	5	0	0	0	0	1	8	15	10
1221	Myristicaceae	*Knema percoriacea* J.Sinclair f. *sarawakensis* W.J.de Wilde	2	30.9	25.96	2	32	1	0	9	18	5	1	1	29	3
224	Dilleniaceae	*Saurauia acuminata* Merr.	2	30.8	22.79	24	36	1	0	0	3	0	0	1	32	3
40	Annonaceae	*Polyalthia cauliflora* Hook.f. & Thomson var. *cauliflora*	2	30.7	26.55	15	31	17	0	4	9	2	0	4	13	4
1297	Moraceae	*Ficus rubrocuspidata* Corner	2	30.6	26.28	14	33	3	0	7	13	0	1	2	28	0
1290	Moraceae	*Ficus oleifolia* King subsp. *oleifolia*	2	30.3	22.42	23	34	19	0	0	11	0	1	0	10	2
371	Flacourtiaceae	*Hydnocarpus woodii* Merr.	2	30.2	23.08	8	34	1	0	2	17	1	0	5	28	4
1159	Loranthaceae	*Barathranthus axanthus* Korth. Miq.	2	29.9	22.57	2	32	10	4	11	14	2	2	10	14	0
97	Euphorbiaceae	*Aporosa sarawakensis* Schot	2	29.8	27.45	10	31	12	1	7	12	2	9	1	16	0
410	Scyphostegiaceae	*Scyphostegia borneensis* Stapf	2	29.8	21.52	29	36	5	0	0	1	0	0	5	20	4
76	Euphorbiaceae	*Antidesma pendulum* Hook.f.	2	29.3	21.15	0	36	0	0	0	12	1	19	9	22	1
890	Liliaceae	*Smilax odoratissima* Blume	2	29	20.73	21	36	2	0	1	12	4	4	0	11	9
1036	Sapindaceae	*Nephelium daedaleum* Radlk.	2	28.9	30.14	1	29	1	0	12	15	3	12	4	21	2
790	Dipterocarpaceae	*Shorea fallax* Meijer	2	28.2	31.49	16	30	1	0	3	0	0	8	4	27	11
482	Ericaceae	*Vaccinium clementis* Merr.	2	27.9	23.85	27	33	11	0	1	9	1	0	1	13	5
777	Dipterocarpaceae	*Shorea biawak* P.S.Ashton	2	27.8	16.58	0	37	1	2	7	8	10	6	3	24	3
954	Polygalaceae	*Xanthophyllum subcoriaceum* Chodat Meijden	2	27.4	27.13	21	31	4	0	0	6	3	7	3	17	8
1403	Orchidaceae	*Dendrochilum pallidiflavens* Blume var. *pallidiflavens*	2	27.4	18.22	32	38	3	0	2	3	0	1	2	8	12
108	Euphorbiaceae	*Baccaurea pubera* Miq. Müll.Arg.	2	27.2	25.71	13	32	4	0	2	7	1	10	5	22	5
826	Dipterocarpaceae	*Shorea patoiensis* P.S.Ashton	2	27	26.76	14	31	0	0	0	2	8	3	1	18	22
187	Euphorbiaceae	*Trigonopleura malayana* Hook.f.	2	26.8	35.27	13	28	1	0	3	4	0	11	6	26	9
456	Ericaceae	*Rhododendron brookeanum* Lowe & Lindl. var. *gracile* Low ex Lindl. Sleumer	2	26.7	32.58	11	27	11	0	0	13	2	3	5	20	9
317	Burseraceae	*Santiria apiculata* A.W.Benn. var. *apiculata*	2	26.5	27.21	14	31	1	0	0	6	0	3	6	30	9
1086	Meliaceae	*Dysoxylum cauliflorum* Hiern	2	26.5	30.72	16	29	4	0	1	8	1	8	1	28	3
132	Euphorbiaceae	*Macaranga brevipetiolata* Airy Shaw	2	26.4	34.67	13	27	3	0	9	15	1	9	0	20	2
232	Dilleniaceae	*Saurauia myrmecoidea* Merr.	2	26.4	11.79	2	41	9	0	0	31	2	0	1	15	0
871	Araliaceae	*Osmoxylon borneense* Seem.	2	26.3	33.3	17	27	14	0	0	14	2	6	0	15	4
562	Fagaceae	*Lithocarpus coopertus* Blanco Rehder	2	26.2	35.61	18	26	9	0	9	7	0	2	3	23	2
228	Dilleniaceae	*Saurauia glabra* Merr.	2	26.1	32.18	11	26	14	0	7	18	3	1	2	18	1
179	Euphorbiaceae	*Pimelodendron griffithianum* Müll. Arg. Benth.	2	25.9	33.73	16	27	2	0	3	4	5	2	4	26	11
311	Burseraceae	*Dacryodes rostrata* Blume H.J.Lam f. *cuspidata* H.J.Lam	2	25.8	22.08	10	33	2	0	0	0	0	7	8	29	11
540	Violaceae	*Rinorea bengalensis* Wall. Kuntze	2	25.8	19.32	10	35	0	0	0	10	0	0	3	35	6
576	Fagaceae	*Lithocarpus pusillus* Soepadmo	2	25.8	21.21	20	34	1	0	9	0	1	3	4	27	2
169	Euphorbiaceae	*Mallotus wrayi* King ex Hook.f.	2	25.3	37.76	17	25	8	0	0	7	2	6	5	23	7
1248	Moraceae	*Artocarpus odoratissimus* Blanco	2	25.2	37.57	10	25	7	0	0	2	4	6	8	21	17
531	Ochnaceae	*Sauvagesia serrata* Korth. Sastre	2	24.9	32.9	5	25	13	4	1	19	3	8	2	18	2
233	Dilleniaceae	*Saurauia planchonii* Hook.f.	2	24.5	13.2	0	39	0	0	0	19	0	0	11	31	0
793	Dipterocarpaceae	*Shorea foxworthyi* Symington	2	24.5	9.12	0	48	0	0	0	6	3	5	0	38	0
268	Thymelaeaceae	*Gonystylus forbesii* Gilg	2	24.3	30.45	10	27	4	1	2	1	6	1	8	25	15
432	Celastraceae	*Perrottetia alpestris* Blume Loes. subsp. *philippinensis* Vidal Ding Hou	2	24.3	16.58	30	36	7	0	0	18	0	0	0	9	0
963	Sabiaceae	*Meliosma sumatrana* Jack Walp.	2	24.2	38.53	20	25	6	0	1	6	2	9	3	24	5

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
733	Dipterocarpaceae	*Dipterocarpus rigidus* Ridl.	2	24.1	12.71	0	38	0	12	8	4	2	0	4	27	4
754	Dipterocarpaceae	*Hopea latifolia* Symington	2	24.1	9.08	0	46	0	0	16	4	0	1	3	30	0
834	Dipterocarpaceae	*Shorea scaberrima* Burck	2	24.1	22.08	4	31	1	2	8	18	6	2	4	26	0
478	Ericaceae	*Vaccinium bancanum* Miq.	2	24	25.03	16	30	1	3	5	8	1	4	1	25	6
751	Dipterocarpaceae	*Hopea fluvialis* P.S.Ashton	2	23.8	12.24	13	42	0	0	0	0	0	2	2	38	3
565	Fagaceae	*Lithocarpus ewyckii* Korth. Rehder	2	23.7	40.03	17	24	7	0	0	20	1	1	1	21	9
730	Dipterocarpaceae	*Dipterocarpus oblongifolius* Blume	2	23.7	39.77	17	24	7	0	3	8	4	8	5	19	6
92	Euphorbiaceae	*Aporosa lunata* Miq. Kurz	2	23.3	27.7	16	28	0	0	19	0	0	1	1	22	12
362	Flacourtiaceae	*Hydnocarpus anomalus* Merr. Sleumer	2	22.9	28.93	23	28	3	0	0	0	1	2	8	20	15
787	Dipterocarpaceae	*Shorea faguetiana* Heim	2	22.8	39.69	8	23	5	0	4	5	8	6	5	22	14
370	Flacourtiaceae	*Hydnocarpus tenuipetala* Sleumer	2	22.7	2.5	0	76	0	0	0	0	0	0	0	23	0
55	Annonaceae	*Popowia pisocarpa* Endl.	2	22.6	40.08	14	24	0	0	1	0	0	10	10	22	17
857	Dipterocarpaceae	*Vatica oblongifolia* Hook.f. subsp. *crassilobata* P.S.Ashton	2	22.6	17.16	21	35	4	0	11	0	1	1	0	20	6
1039	Sapindaceae	*Nephelium subfalcatum* Radlk.	2	22.6	3.85	0	65	0	0	3	0	1	1	0	29	0
1059	Meliaceae	*Aglaia macrocarpa* Miq. Pannell	2	22.6	27.81	15	28	2	0	5	0	1	9	5	15	21
1254	Moraceae	*Ficus androchaete* Corner	2	22.6	32.44	18	26	4	0	0	21	1	5	0	18	6
109	Euphorbiaceae	*Baccaurea sarawakensis* Pax & K.Hoffm.	2	22.5	40.52	9	23	9	0	8	14	3	8	4	21	2
99	Euphorbiaceae	*Aporosa subcaudata* Merr.	2	22.3	41.81	15	24	1	0	0	4	1	9	4	24	18
4	Annonaceae	*Cyathocalyx havilandii* Boerl.	2	22	38.53	3	22	7	2	0	12	15	4	8	18	10
537	Oxalidaceae	*Sarcotheca ochracea* Hallier f.	2	22	11.67	0	39	0	0	5	7	7	0	6	26	10
737	Dipterocarpaceae	*Dipterocarpus verrucosus* Foxw. ex Slooten	2	22	20.31	5	31	1	0	19	3	3	0	3	26	8
155	Euphorbiaceae	*Mallotus caudatus* Merr.	2	21.9	39.75	16	23	4	1	0	1	5	3	7	22	16
434	Celastraceae	*Salacia korthalsiana* Miq.	2	21.9	14.49	5	38	0	0	0	0	0	8	4	27	19
1242	Moraceae	*Artocarpus brevipedunculatus* Jarrett C.C.Berg	2	21.9	36.62	14	24	9	0	0	13	4	0	4	22	9
154	Euphorbiaceae	*Macaranga winkleri* Pax & K.Hoffm.	2	21.6	44	13	22	9	0	2	2	6	8	6	18	14
464	Ericaceae	*Rhododendron longiflorum* Lindl.	2	21.5	30.16	21	26	4	0	16	5	2	1	0	18	7
919	Magnoliaceae	*Magnolia liliifera* L. Baill. var. *liliifera*	2	21.5	30.16	22	27	8	0	0	1	0	8	1	24	9
1328	Orchidaceae	*Arundina graminifolia* D.Don Hochr.	2	21.3	21.1	13	30	10	0	1	17	1	9	2	15	2
28	Annonaceae	*Monocarpia kalimantanensis* P.Kessler	2	21	30.19	7	26	0	5	3	1	2	6	9	25	17
131	Euphorbiaceae	*Macaranga beccariana* Merr.	2	20.9	35.47	7	23	7	0	3	16	5	6	5	23	5
1353	Orchidaceae	*Bulbophyllum odoratum* Blume Lindl.	2	20.9	29.35	24	26	14	0	0	4	3	3	1	15	9
571	Fagaceae	*Lithocarpus leptogyne* Korth. Soepadmo	2	20.8	43.22	15	22	1	1	3	10	3	1	13	21	12
755	Dipterocarpaceae	*Hopea mesuoides* P.S.Ashton	2	20.7	15.45	3	35	1	0	0	0	4	6	2	32	16
70	Euphorbiaceae	*Antidesma leucopodum* Miq.	2	20.4	48.11	15	21	3	0	10	9	2	8	4	20	10
297	Burseraceae	*Canarium littorale* Blume f. rufum Leenh.	2	20	25.92	6	25	3	11	0	3	13	2	14	20	4
948	Polygalaceae	*Xanthophyllum pulchrum* King	2	19.8	8.56	0	45	0	0	0	2	0	2	4	36	10
1098	Meliaceae	*Sandoricum borneense* Miq.	2	19.2	42.54	6	20	10	0	14	17	5	5	7	14	2
1045	Sapindaceae	*Xerospermum noronhianum* Blume Blume	2	19	48.55	14	19	11	7	5	12	1	6	2	18	6
999	Rosaceae	*Prunus arborea* Blume Kalkman var. *densa* King Kalkman	2	18.3	31.82	21	23	17	0	0	11	1	1	3	11	10
1299	Moraceae	*Ficus schwarzii* Koord.	2	18.3	47.46	11	19	11	0	4	18	7	7	4	16	3
387	Proteaceae	*Helicia attenuata* Jack Blume	2	18	33.78	17	23	3	0	5	15	0	8	1	20	7
776	Dipterocarpaceae	*Shorea balanocarpoides* Symington	2	18	6.04	0	49	0	0	9	0	0	0	1	41	0
753	Dipterocarpaceae	*Hopea kerangasensis* P.S.Ashton	2	17.6	13.4	0	33	1	1	17	6	6	4	9	23	0
630	Cornaceae	*Mastixia rostrata* Blume subsp. *caudatifolia* Merr.	2	17.2	38.41	21	22	2	0	0	5	0	7	6	16	21
798	Dipterocarpaceae	*Shorea hemsleyana* King ex Foxw. subsp. *grandiflora* Brandis P.S.Ashton	2	16.9	13.8	0	31	0	1	15	21	9	2	7	13	0
310	Burseraceae	*Dacryodes nervosa* H.J.Lam Leenh.	2	16.8	18.04	0	27	0	7	6	7	11	1	14	21	5
781	Dipterocarpaceae	*Shorea coriacea* Burck	2	15.7	16.75	21	30	11	0	0	2	1	11	4	13	7
816	Dipterocarpaceae	*Shorea ochracea* Symington	2	14.7	37.95	4	18	0	8	8	9	12	2	7	18	13
1184	Viscaceae	*Ginalloa arnottiana* Korth.	2	14	15.86	19	30	0	0	0	7	1	0	0	29	14
721	Dipterocarpaceae	*Dipterocarpus crinitus* Dyer	2	13.7	24.85	0	23	1	0	12	2	12	9	5	19	18
765	Dipterocarpaceae	*Hopea vesquei* Heim	2	13.1	28.12	0	20	2	12	3	11	16	9	10	17	1
1214	Myristicaceae	*Knema lunduensis* J.Sinclair W.J.de Wilde	2	12.8	23.69	2	22	0	1	14	18	1	14	6	22	2
748	Dipterocarpaceae	*Hopea coriacea* Burck	2	12.4	13.89	0	27	0	1	15	7	11	2	9	26	0

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
909	Araucariaceae	*Agathis endertii* Meijer Drees	2	11.8	16.78	6	25	5	0	0	1	23	1	4	15	19
360	Flacourtiaceae	*Homalium moultonii* Merr.	2	4.8	2.38	0	41	0	0	15	0	0	0	10	34	0
1007	Rosaceae	*Rubus benguetensis* Elmer	3	36.8	19.86	35	16	40	0	0	3	1	2	0	3	1
1357	Orchidaceae	*Bulbophyllum sopoetanense* Schltr.	3	36.8	20.31	34	10	41	0	0	7	1	3	0	4	2
1428	Orchidaceae	*Pholidota ventricosa Blume* Rchb.f.	3	33.8	16.88	38	15	42	0	0	3	0	0	0	2	0
513	Clethraceae	*Clethra canescens* Reinw. ex Blume var. *clementis* Merr. Sleumer	3	25.9	26.4	30	9	31	0	0	10	1	0	0	14	5
1344	Orchidaceae	*Bulbophyllum lobbii* Lindl.	3	25.9	18.13	24	32	33	0	0	3	0	0	1	6	1
1391	Orchidaceae	*Dendrobium indivisum* Blume Miq.	3	24.3	36.52	21	13	25	0	0	5	2	8	0	19	6
892	Olacaceae	*Erythropalum scandens* Blume	3	18.6	39.17	21	0	22	0	0	15	5	4	0	20	12
1285	Moraceae	*Ficus megaleia* Corner	3	15.4	36.1	16	7	20	0	6	9	1	9	11	9	11
927	Podocarpaceae	*Nageia wallichiana* Kuntze	3	10.2	19.4	14	1	22	0	0	17	6	5	18	9	9
720	Dipterocarpaceae	*Dipterocarpus costulatus* Sloot.	4	33.4	12.62	0	0	0	50	0	0	0	8	1	36	6
118	Euphorbiaceae	*Chondrostylis bancana* Boerl.	4	27.4	13.4	0	3	0	42	0	14	12	17	0	10	3
714	Dipterocarpaceae	*Dipterocarpus borneensis* Slooten	4	27.3	19.66	0	5	0	34	0	5	11	6	11	25	3
1394	Orchidaceae	*Dendrobium lobatum* Blume Miq.	4	25.3	16.11	0	0	0	40	0	0	0	27	0	15	18
303	Burseraceae	*Canarium patentinervium* Miq.	4	18.6	29.53	0	5	0	23	0	2	13	2	23	23	8
1162	Loranthaceae	*Dendrophthoe longituba* Elmer Danser	4	15.8	12.77	8	20	0	32	4	1	1	0	7	23	5
1073	Meliaceae	*Chisocheton amabilis* Miq.	5	61.2	14.36	0	0	0	0	63	0	0	9	16	10	3
1278	Moraceae	*Ficus heterophylla* L.f.	5	50.6	14.97	0	0	0	0	55	0	1	1	24	6	14
1167	Loranthaceae	*Helixanthera setigera* Korth. Danser	5	45.9	12.42	0	0	0	0	57	0	0	0	29	10	5
770	Dipterocarpaceae	*Shorea albida* Symington	5	42.8	8.7	0	0	0	0	69	0	0	4	1	25	1
438	Ericaceae	*Costera ovalifolia* J.J.Sm.	5	42.2	17.42	13	8	0	0	49	0	2	3	3	13	10
189	Amaranthaceae	*Alternanthera sessilis* L. R.Br. ex DC.	5	41	20.24	0	0	0	0	42	2	0	3	41	7	5
1116	Caesalpiniaceae	*Bauhinia kockiana* Korth. var. *velutina* de Wit K.Larsen & S.S.Larsen	5	35	17.7	0	1	0	0	46	12	0	32	1	8	0
1182	Loranthaceae	*Scurrula lepidota* Blume G.Don	5	34.2	19.02	0	0	0	0	40	1	0	3	38	6	13
950	Polygalaceae	*Xanthophyllum ramiflorum* Meijden	5	32.5	20.3	0	2	0	34	36	2	5	1	7	14	0
740	Dipterocarpaceae	*Dryobalanops fusca* Slooten	5	31.9	12.62	0	0	0	0	47	4	1	2	39	7	0
1267	Moraceae	*Ficus depressa* Blume	5	30.5	24.97	12	2	0	0	37	2	0	15	1	27	4
843	Dipterocarpaceae	*Shorea teysmanniana* Dyer	5	27.6	16.03	0	0	0	1	40	2	2	4	30	19	2
95	Euphorbiaceae	*Aporosa octandra* Buch.-Ham. ex D.Don Vickery var. *malesiana* Schot	5	27.3	18.97	0	0	0	1	38	9	2	18	19	11	3
1416	Orchidaceae	*Eria xanthocheila* Ridl.	5	26.8	20.16	1	8	5	0	35	21	4	15	4	6	0
1044	Sapindaceae	*Xerospermum laevigatum* Radlk. subsp. *acuminatum* Radlk. Leenh.	5	25.2	21.23	0	0	0	7	34	0	6	0	14	27	11
148	Euphorbiaceae	*Macaranga puncticulata* Gage	5	24.8	30.43	1	0	0	11	28	1	2	11	22	19	5
181	Euphorbiaceae	*Sauropus androgynus* L. Merr.	5	24.2	34.7	0	5	0	0	26	9	5	3	19	22	9
1311	Moraceae	*Ficus villosa* Blume	5	23.4	30.21	19	4	8	0	28	5	0	1	3	19	12
1314	Moraceae	*Parartocarpus venenosus* Becc.	5	22.3	25.52	15	4	2	4	31	0	2	8	1	23	9
267	Thymelaeaceae	*Gonystylus bancanus* Miq. Kurz	5	22.1	20.35	0	1	0	0	32	11	0	0	26	28	1
1128	Caesalpiniaceae	*Copaifera palustris* Symington de Wit	5	21.9	31.21	0	5	0	5	26	5	9	9	20	20	3
1179	Loranthaceae	*Macrosolen retusus* Jack Miq.	5	21.6	20.4	0	0	1	11	31	2	4	12	26	12	0
817	Dipterocarpaceae	*Shorea ovalis* Korth. Blume	5	21.4	43.85	3	1	0	1	22	4	9	4	19	20	16
1097	Meliaceae	*Sandoricum beccarianum* Baill.	5	21	19.24	0	0	0	3	32	0	1	1	31	23	9
413	Campanulaceae	*Pentaphragma acuminatum* Airy Shaw	5	20.9	17.24	32	0	25	0	34	2	0	0	0	6	0
403	Rhizophoraceae	*Combretocarpus rotundatus* Miq. Danser	5	20.1	25.79	0	0	0	11	27	1	1	4	21	24	12
783	Dipterocarpaceae	*Shorea cuspidata* P.S.Ashton	5	20	7.01	0	4	0	1	52	4	0	2	9	28	0
343	Dichapetalaceae	*Dichapetalum gelonioides* Roxb. Engl. subsp. *sumatranum* Miq. Leenh.	5	19.6	23.31	10	9	0	0	29	0	0	1	9	24	17
338	Connaraceae	*Ellipanthus tomentosus* Kurz var. *tomentosus*	5	18.1	11.09	0	0	0	3	40	1	1	3	15	36	0
1191	Myristicaceae	*Horsfieldia crassifolia* Hook.f. & Thomson Warb.	5	17.2	11.01	0	0	0	0	39	0	0	0	19	33	8
628	Cornaceae	*Mastixia eugenioides* K.M.Matthew	5	16.9	27.2	0	18	0	1	23	9	11	2	16	20	1
1185	Viscaceae	*Viscum ovalifolium* DC.	5	16.8	35.82	7	4	2	0	22	15	0	11	14	19	5
595	Cyperaceae	*Fimbristylis ferruginea* L. Vahl	5	16.7	14.1	0	0	0	2	34	0	0	6	32	22	4
3	Annonaceae	*Cyathocalyx biovulatus* Boerl.	5	16.1	7.79	0	0	0	0	45	0	0	1	17	37	0
880	Araliaceae	*Schefflera tetrandra* Merr.	5	15.9	16.47	0	20	0	18	28	6	3	1	4	20	0
1266	Moraceae	*Ficus deltoidea* Jack subsp. *motleyana* Miq. C.C.Berg	5	14.2	43.79	3	6	0	13	17	12	5	9	12	17	4
1156	Daphniphyllaceae	*Daphniphyllum laurinum* Benth. Baill.	5	14.1	13	8	4	0	0	34	1	0	6	4	33	11

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
119	Euphorbiaceae	*Chondrostylis kunstleri* King ex Hook.f. Airy Shaw	5	13.7	6.55	0	36	0	6	39	1	0	1	3	14	0
239	Dioscoriaceae	*Dioscorea pyrifolia* Kunth	5	13.6	18.83	6	7	0	0	27	4	1	2	20	19	13
326	Burseraceae	*Santiria rubiginosa* Blume var. *pedicellata* Ridl. Kalkman	5	13.3	11.5	0	0	0	12	32	2	0	3	28	22	0
1409	Orchidaceae	*Eria leiophylla* Lindl.	5	12.2	26.08	21	3	5	0	23	14	0	7	0	22	6
116	Euphorbiaceae	*Cephalomappa paludicola* Airy Shaw	5	10	1.67	0	0	0	0	72	0	0	3	21	4	0
693	Anacardiaceae	*Parishia insignis* Hook.f.	5	9.5	14.49	0	6	0	0	27	0	0	17	5	25	20
1006	Rosaceae	*Prunus turfosa* Kalkman	5	8.9	7.9	0	4	0	0	33	23	2	13	7	18	0
261	Thymelaeaceae	*Aetoxylon sympetalum* Steenis & Domke Airy Shaw	5	8.5	26.21	0	10	0	11	17	9	6	13	17	16	0
102	Euphorbiaceae	*Baccaurea brevipes* Hook.f.	5	3.9	2.94	0	27	0	0	34	2	0	7	13	17	0
573	Fagaceae	*Lithocarpus meijeri* Soepadmo	6	26.6	19.06	13	22	7	0	0	36	0	0	0	21	0
138	Euphorbiaceae	*Macaranga hosei* King ex Hook.f.	6	22.6	31.4	1	11	2	0	13	26	7	12	5	22	2
840	Dipterocarpaceae	*Shorea stenoptera* Burck	6	22	27.46	0	20	10	2	3	25	9	3	19	10	0
135	Euphorbiaceae	*Macaranga costulata* Pax & K.Hoffm.	6	20.5	36.87	14	12	8	0	4	23	1	6	10	21	1
230	Dilleniaceae	*Saurauia javanica* Nees Hoogland	6	17.7	39.23	19	6	18	0	0	21	1	0	9	19	5
1277	Moraceae	*Ficus hemsleyana* King	6	15.9	38.18	14	19	12	0	3	20	2	8	2	20	1
536	Oxalidaceae	*Sarcotheca macrophylla* Blume	6	12.2	38.87	0	14	8	10	8	16	11	13	6	6	8
234	Dilleniaceae	*Saurauia subcordata* Korth.	6	10.6	9.79	19	0	12	0	0	33	0	0	18	14	4
785	Dipterocarpaceae	*Shorea elliptica* Burck	6	10.4	14.82	0	24	5	1	12	24	3	9	7	16	0
884	Liliaceae	*Peliosanthes teta* Andrews subsp. *humilis* Andr. Jessop	8	36.2	17.09	3	0	0	0	1	3	0	52	12	28	1
1088	Meliaceae	*Dysoxylum grande* Hiern	8	21	17.72	20	0	1	0	0	3	2	40	0	33	1
52	Annonaceae	*Polyalthia xanthopetala* Merr.	8	16	16.16	1	0	0	0	0	1	0	34	23	33	8
985	Mimosaceae	*Archidendron clypearia* Jack I.C.Nielsen var. *casai* Blanco I.C.Nielsen	8	15.5	38.89	12	0	3	0	0	9	10	21	15	16	15
626	Bignoniaceae	*Radermachera gigantea* Blume Miq.	8	14.4	5.82	0	0	0	0	6	0	0	56	3	6	28
690	Anacardiaceae	*Melanochyla beccariana* Oliv.	8	7.3	8.59	21	0	9	0	0	2	0	33	0	31	4
599	Cyperaceae	*Fuirena umbellata* Rottb.	9	52.6	15.56	0	0	0	0	4	4	3	1	56	14	19
210	Convolvulaceae	*Ipomoea mauritiana* Jacq.	9	52.2	18.78	0	0	0	0	0	4	1	7	52	22	13
898	Opiliaceae	*Lepionurus sylvestris* Blume	9	43.2	21.2	0	0	0	0	0	5	2	6	45	31	9
593	Cyperaceae	*Cyperus polystachyos* Rottb.	9	38.9	12.94	0	0	0	3	16	0	1	7	53	14	6
604	Cyperaceae	*Rhynchospora corymbosa* L. Britton	9	37.6	21.52	1	0	0	3	17	1	1	1	41	17	19
409	Rhizophoraceae	*Rhizophora mucronata* Lam.	9	36.7	14.34	0	0	0	0	3	0	0	2	50	27	16
596	Cyperaceae	*Fimbristylis globulosa* Retz. Kunth	9	34.3	20.72	0	0	0	0	19	1	2	3	40	24	11
523	Icacinaceae	*Phytocrene anomala* Merr.	9	34.1	17.71	0	0	0	0	0	1	0	16	45	28	10
606	Cyperaceae	*Scirpodendron ghaeri* Gaertn. Merr.	9	32.1	15.46	0	0	0	0	0	0	1	11	46	15	27
592	Cyperaceae	*Cyperus kyllingia* Endl.	9	31.7	27.43	18	0	0	0	0	1	4	2	35	17	23
636	Onagraceae	*Ludwigia hyssopifolia* Don Exell	9	31.7	30.61	10	0	0	0	0	3	2	2	33	28	22
775	Dipterocarpaceae	*Shorea balangeran* Korth. Burck	9	30.9	19.31	0	0	0	5	32	2	4	5	37	4	12
287	Burseraceae	*Canarium caudatum* King f. *auriculiferum* Leenh.	9	30	17.36	0	0	0	1	0	2	3	10	42	29	14
634	Lentibulariaceae	*Utricularia aurea* Lour.	9	28.8	13.09	0	0	0	5	33	0	0	3	44	6	8
1217	Myristicaceae	*Knema mogeana* W.J.de Wilde	9	28.7	27.13	0	0	0	0	0	16	2	20	33	21	8
598	Cyperaceae	*Fimbristylis pauciflora* R.Br.	9	28.6	13.85	0	11	0	3	15	1	0	1	43	22	4
379	Flacourtiaceae	*Scolopia macrophylla* Wight & Arn. Chois.	9	27.3	11.66	1	0	0	0	0	0	0	7	49	32	10
629	Cornaceae	*Mastixia pentandra* Blume subsp. *scortechinii* King Matthew	9	27.1	28.2	16	0	1	0	2	3	9	3	32	8	25
406	Rhizophoraceae	*Pellacalyx axillaris* Korth.	9	26.8	24.47	8	2	2	0	2	5	6	0	33	29	12
761	Dipterocarpaceae	*Hopea sangal* Korth.	9	26.2	37.92	1	0	0	0	0	6	6	16	27	24	19
1229	Myristicaceae	*Myristica elliptica* Hook.f. & Thomson var. *elliptica* Hook.f. & Thomson	9	26	24.48	0	10	0	0	0	17	4	5	32	26	6
1104	Meliaceae	*Xylocarpus granatum* Koenig	9	25.6	31.93	0	1	0	6	17	3	5	7	28	23	10
994	Mimosaceae	*Mimosa pigra* L.	9	24.8	15.62	0	0	0	0	33	0	0	3	38	5	21
973	Mimosaceae	*Acacia borneensis* I.C.Nielsen	9	24.5	13.03	0	0	0	0	0	0	0	12	44	11	33
396	Rhizophoraceae	*Bruguiera gymnorhiza* L. Lam.	9	23.8	13.71	0	0	0	0	19	0	0	10	42	24	4
101	Euphorbiaceae	*Baccaurea bracteata* Müll.Arg.	9	23.7	38.81	1	1	0	7	20	3	7	3	24	20	14
165	Euphorbiaceae	*Mallotus muticus* (Müll.Arg.) Airy Shaw	9	22.9	42.96	9	0	0	7	5	2	7	5	24	22	20
274	Thymelaeaceae	*Gonystylus velutinus* Airy Shaw	9	22.3	40.53	0	4	0	0	1	8	18	10	23	17	18
1068	Meliaceae	*Aglaia simplicifolia* Bedd. Harms	9	22.3	20.96	5	0	0	0	0	7	3	11	34	29	12
325	Burseraceae	*Santiria rubiginosa* Blume	9	22.2	23.07	0	0	0	1	29	2	2	1	30	21	14

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
68	Euphorbiaceae	*Antidesma coriaceum* Tul.	9	21.9	26.94	2	5	0	11	20	0	1	1	27	23	9
183	Euphorbiaceae	*Shirakiopsis indica* Willd. Esser	9	20.9	14.12	0	0	0	5	3	0	0	0	39	38	15
429	Celastraceae	*Lophopetalum wightianum* Arn.	9	20.1	16	0	0	0	2	0	0	1	16	37	12	32
989	Mimosaceae	*Archidendron jiringa* Jack I.C.Nielsen	9	20.1	20.15	0	0	0	0	0	1	12	12	32	16	28
197	Combretaceae	*Combretum tetralophum* C.B.Clarke	9	19.6	11.01	0	0	0	0	11	1	1	0	42	22	23
708	Labiatae	*Hyptis capitata* Jacq.	9	19	23.71	8	0	1	1	12	3	0	4	29	28	15
861	Dipterocarpaceae	*Vatica rassak* Korth. Blume	9	19	44.77	5	0	0	5	20	1	10	6	20	16	17
1146	Caesalpiniaceae	*Sindora coriacea* Prain	9	18.9	14.77	6	0	0	0	0	0	2	9	37	21	26
1310	Moraceae	*Ficus urnigera* Miq.	9	18.3	39.37	22	1	4	0	10	4	4	1	22	12	19
178	Euphorbiaceae	*Omphalea sargentii* Merr.	9	17.9	21.76	9	0	0	0	0	0	0	13	30	25	22
688	Anacardiaceae	*Mangifera laurina* Blume	9	17.2	19.2	0	0	0	5	0	1	3	4	31	26	30
591	Cyperaceae	*Cyperus javanicus* Houtt.	9	16.9	12.23	0	0	0	11	13	2	1	2	36	25	11
218	Dilleniaceae	*Dillenia eximia* Miq.	9	16.8	16.7	0	0	0	0	0	8	3	1	32	27	29
1161	Loranthaceae	*Dendrophthoe curvata* Blume Miq.	9	16.7	18.3	1	0	0	6	24	3	1	4	30	25	5
667	Ulmaceae	*Trema cannabina* Lour.	9	16.6	60.36	2	1	0	5	11	8	10	14	17	16	16
896	Olacaceae	*Strombosia javanica* Blume	9	14.4	10.35	0	0	0	0	18	0	1	9	37	5	30
199	Combretaceae	*Quisqualis indica* L.	9	13.9	22.89	0	0	0	8	17	1	10	10	24	10	21
594	Cyperaceae	*Fimbristylis cymosa* R.Br.	9	13.8	14.62	0	0	0	5	23	1	0	3	30	27	10
345	Dichapetalaceae	*Dichapetalum grandifolium* Ridl.	9	13.6	15.88	15	0	0	0	0	0	1	9	31	21	22
589	Cyperaceae	*Cyperus brevifolius* Rottb. Hassk.	9	12.9	24.94	16	5	1	2	1	1	3	9	24	24	14
1275	Moraceae	*Ficus grossularioides* Burm.f. var. *grossularioides*	9	11.3	17.17	0	3	0	0	22	18	6	5	25	15	7
605	Cyperaceae	*Schoenus calostachyus* R.Br. Poir.	9	11.2	12.08	0	0	0	2	28	1	2	9	30	26	2
707	Labiatae	*Hyptis brevipes* Poit.	9	11.1	35.19	12	0	9	3	14	4	7	7	18	12	15
642	Symplocaceae	*Symplocos cochinchinensis* Lour. S.Moore subsp. *cochinchinensis*	9	10.6	20.68	19	0	15	0	14	8	1	11	22	3	7
709	Labiatae	*Leucas zeylanica* L. R.Br.	9	8	8.32	0	0	0	0	27	0	0	4	31	30	8
66	Euphorbiaceae	*Aleurites moluccana* L. Willd.	9	6.4	7.99	6	0	0	6	6	0	0	5	29	26	22
956	Polygalaceae	*Xanthophyllum trichocladum* Chodat	10	55.6	13.51	1	3	0	0	0	4	0	2	11	70	10
273	Thymelaeaceae	*Gonystylus stenosepalus* Airy Shaw	10	55	13.1	1	15	0	0	0	15	0	1	1	68	0
63	Annonaceae	*Uvaria sorzogonensis* C.Presl	10	53.3	9.75	4	0	0	0	0	0	0	1	1	83	11
209	Convolvulaceae	*Erycibe stenophylla* Hoogland	10	51.1	11.09	8	1	0	0	0	0	0	1	0	76	14
768	Dipterocarpaceae	*Shorea agami* P.S.Ashton	10	50.4	11.12	5	1	0	0	0	0	0	0	4	75	15
1175	Loranthaceae	*Loxanthera speciosa* Blume	10	50.4	18.78	4	3	0	2	12	1	0	6	15	55	1
1228	Myristicaceae	*Myristica depressa* W.J.de Wilde	10	50.4	11.67	0	1	0	1	14	0	0	1	13	70	1
967	Chrysobalaceae	*Kostermanthus heteropetalus* Scort. ex King Prance	10	50.3	11.43	11	0	0	0	0	0	0	0	0	75	14
93	Euphorbiaceae	*Aporosa nigricans* Hook.f.	10	50.1	8.71	3	0	0	0	0	0	0	0	1	85	11
767	Dipterocarpaceae	*Shorea acuminatissima* Symington	10	49.8	12.23	4	0	0	0	0	0	0	0	7	71	18
797	Dipterocarpaceae	*Shorea havilandi* Brandis	10	49.8	19.81	4	2	0	0	13	3	0	6	10	54	8
853	Dipterocarpaceae	*Vatica mangachapoi* Blanco	10	48.9	15.38	8	0	0	0	13	0	0	2	0	62	16
500	Loganiaceae	*Norrisia maior* Soler.	10	48	16.65	12	3	0	0	8	0	0	0	0	59	18
47	Annonaceae	*Polyalthia insignis* Hook.f. Airy Shaw	10	46.7	14.22	17	0	0	0	1	0	0	1	5	64	11
1125	Caesalpiniaceae	*Caesalpinia major* Medik. Dandy & Exell	10	46.5	10.1	0	0	0	0	5	0	0	2	9	74	9
1120	Caesalpiniaceae	*Bauhinia wrayi* Prain ex King var. *borneensis* K.Larsen & S.S.Larsen	10	46	20.88	3	9	0	0	0	7	2	8	9	50	13
1224	Myristicaceae	*Knema stylosa* W.J.de Wilde W.J.de Wilde	10	45.9	16.15	19	6	0	0	0	0	0	5	1	59	9
288	Burseraceae	*Canarium caudatum* King f. *caudatum*	10	45.7	16.14	11	10	0	0	2	0	4	0	6	57	9
836	Dipterocarpaceae	*Shorea scrobiculata* Burck	10	45	15.82	10	26	0	0	0	0	0	1	4	55	5
1118	Caesalpiniaceae	*Bauhinia semibifida* Roxb.	10	44.8	17.37	8	7	0	0	1	0	1	17	0	56	9
366	Flacourtiaceae	*Hydnocarpus gracilis* Slooten Sleumer	10	44.6	12.03	5	0	0	0	0	0	2	2	0	68	23
936	Polygalaceae	*Xanthophyllum adenotus* Miq. var. *adenotus*	10	44.6	15.33	14	14	0	0	0	0	0	0	6	58	9
1274	Moraceae	*Ficus globosa* Blume	10	44.4	11.72	7	0	0	0	2	0	0	11	8	69	3
1021	Sapindaceae	*Guioa bijuga* Hiern Radlk.	10	44.2	14.43	15	0	0	0	0	0	0	0	2	62	20
543	Violaceae	*Rinorea longiracemosa* Kurz Craib	10	44	21.81	17	1	0	0	0	0	0	8	12	50	13
32	Annonaceae	*Orophea corymbosa* Blume Miq.	10	43.7	12.66	3	0	0	2	0	2	1	6	14	64	9
389	Proteaceae	*Helicia fuscotomentosa* Suess.	10	43.7	12.56	21	2	2	0	0	3	0	1	5	66	1
845	Dipterocarpaceae	*Shorea xanthophylla* Symington	10	43.4	16.1	9	21	0	0	0	1	0	0	4	54	12
805	Dipterocarpaceae	*Shorea laxa* Slooten	10	42.9	15.72	4	5	0	0	3	0	1	5	3	57	21
695	Anacardiaceae	*Parishia sericea* Ridl.	10	42.6	14	17	4	0	0	0	0	0	2	1	62	14

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
373	Flacourtiaceae	*Ryparosa acuminata* Merr.	10	42.4	20.22	12	19	0	0	0	3	0	4	7	48	7
428	Celastraceae	*Lophopetalum subobovatum* King	10	42.4	14.29	2	0	0	20	0	2	0	5	14	56	1
1074	Meliaceae	*Chisocheton erythrocarpus* Hiern	10	42	15.1	7	2	0	19	0	0	1	1	0	56	14
979	Mimosaceae	*Albizia kostermansii* I.C.Nielsen	10	41.9	12.16	6	0	0	0	0	0	0	13	0	66	14
1010	Rosaceae	*Rubus moluccanus* L. var. *angulosus* Kalkman	10	41.6	13.02	7	1	0	0	0	0	0	14	0	64	14
1106	Caesalpiniaceae	*Bauhinia diptera* Miq.	10	41.5	11.27	1	0	0	0	1	0	0	0	11	66	21
1079	Meliaceae	*Chisocheton pentandrus* Blanco Merr. subsp. *paucijugus* Miq. Mabb.	10	41.4	22.34	18	4	0	0	0	0	0	9	6	48	15
498	Loganiaceae	*Fagraea ridleyi* King & Gamble	10	41.3	13.58	15	6	0	0	0	5	0	1	11	59	2
809	Dipterocarpaceae	*Shorea macropterafolia* P.S.Ashton	10	40	14.9	6	29	0	0	0	2	2	0	0	52	9
618	Hypericaceae	*Cratoxylum arborescens* Vahl Blume var. *arborescens*	10	39.9	18.56	16	14	0	0	3	0	0	0	5	49	12
801	Dipterocarpaceae	*Shorea isoptera* P.S.Ashton	10	39.9	7.08	0	2	0	0	0	0	0	0	15	82	1
21	Annonaceae	*Mezzettia havilandii* Boerl. Ridl.	10	39.8	14.49	2	5	0	5	8	1	0	1	22	54	2
222	Dilleniaceae	*Dillenia sumatrana* Miq.	10	39.8	18.98	5	15	0	12	2	1	0	1	7	46	12
157	Euphorbiaceae	*Mallotus floribundus* (Blume) Műll. Arg.	10	39.7	25.06	4	2	0	0	0	10	1	3	27	41	12
780	Dipterocarpaceae	*Shorea confusa* P.S.Ashton	10	39.6	10.7	0	0	0	0	2	0	0	0	15	66	17
838	Dipterocarpaceae	*Shorea slootenii* G.H.S.Wood	10	39.6	9.36	0	16	0	0	0	0	12	0	3	67	1
245	Gnetaceae	*Gnetum diminutum* Markgr.	10	39.4	18.76	8	18	0	6	3	2	0	1	5	47	11
341	Dichapetalaceae	*Dichapetalum gelonioides* Roxb. Engl.	10	39.4	12.51	4	2	0	0	0	8	2	8	0	62	14
355	Flacourtiaceae	*Flacourtia jangomas* Lour. Raeusch.	10	39.4	15.03	2	1	0	7	10	0	0	13	9	55	4
399	Rhizophoraceae	*Carallia borneensis* Oliv.	10	39.4	22.3	9	0	0	0	0	1	0	0	21	45	23
27	Annonaceae	*Monocarpia borneensis* Mols & P.J.A.Kessler	10	39.3	14.05	4	9	0	0	1	0	1	13	0	58	14
339	Connaraceae	*Rourea mimosoides* Vahl Planch. f. *obtusifolia* Leenh.	10	39.1	10.41	4	0	0	1	0	2	0	6	6	68	12
121	Euphorbiaceae	*Euphorbia hirta* L.	10	39	8.98	0	0	0	2	1	0	2	1	18	71	6
1150	Caesalpiniaceae	*Sympetalandra unijuga* Airy Shaw Steenis	10	39	15.08	2	0	0	0	0	8	0	2	14	55	19
26	Annonaceae	*Mitrephora maingayi* Hook.f. & Thomson	10	38.4	16.89	5	5	0	0	24	0	0	3	0	50	11
331	Connaraceae	*Cnestis palala* Lour. Merr. subsp. *diffusa* Blanco Andreas	10	38.3	13.72	7	0	0	25	1	0	0	0	1	55	11
1186	Myristicaceae	*Endocomia rufirachis* J.Sinclair W.J.de Wilde	10	38.3	17.97	3	1	0	0	0	1	0	0	20	49	25
758	Dipterocarpaceae	*Hopea nutans* Ridl.	10	38.2	11.31	0	1	0	0	10	0	0	1	16	62	10
212	Convolvulaceae	*Merremia borneensis* Merr.	10	38	11.36	22	0	0	0	0	0	0	0	3	66	9
315	Burseraceae	*Dacryodes rugosa* Blume H.J.Lam	10	38	18.57	9	2	0	0	0	1	0	9	6	50	23
759	Dipterocarpaceae	*Hopea pentanervia* Symington ex G.H.S.Wood	10	38	5.29	0	0	0	0	0	0	0	0	0	95	5
12	Annonaceae	*Goniothalamus suluensis* Merr.	10	37.6	15.47	1	0	0	0	0	0	1	29	0	56	13
1024	Sapindaceae	*Guioa pubescens* Zoll. & Moritzi Radlk.	10	37.4	17.66	29	2	0	0	0	0	2	0	0	52	15
977	Mimosaceae	*Adenanthera pavonina* L.	10	37.3	21.86	6	4	0	0	0	1	0	16	22	44	6
520	Icacinaceae	*Gonocaryum minus* Sleumer	10	37	13.99	2	31	0	9	3	1	1	1	1	49	1
1031	Sapindaceae	*Lepisanthes senegalensis* Poir. Leenh.	10	37	14.18	3	0	0	0	0	0	0	4	10	56	27
1108	Caesalpiniaceae	*Bauhinia excelsa* Blume ex Miq. Prain var. *megalantha* Merr. K.Larsen & S.S.Larsen	10	37	11.05	10	0	0	4	0	0	0	2	5	64	15
750	Dipterocarpaceae	*Hopea dyeri* Heim	10	36.8	17.51	9	13	0	0	0	11	0	2	9	48	8
784	Dipterocarpaceae	*Shorea domatiosa* P.S.Ashton	10	36.7	4.71	0	0	0	0	0	0	0	0	0	100	0
966	Chrysobalaceae	*Atuna racemosa* Raf.	10	36.4	17.63	6	0	0	0	5	0	0	3	6	50	31
1173	Loranthaceae	*Lepidaria pulchella* Danser	10	36.2	22.29	11	24	1	1	4	2	2	0	3	41	11
176	Euphorbiaceae	*Neoscortechinia sumatrensis* S.Moore	10	36	11.89	3	12	0	0	0	1	0	3	17	57	6
1115	Caesalpiniaceae	*Bauhinia kockiana* Korth. var. *kockiana*	10	36	23.57	22	6	0	0	0	0	0	0	3	43	26
285	Burseraceae	*Canarium asperum* Benth. f. *villosum* Blume H.J.Lam	10	35.9	15.23	12	0	0	2	0	0	0	14	5	54	12
502	Loganiaceae	*Strychnos ovata* A.W.Hill	10	35.9	14.24	2	1	0	0	3	0	0	1	12	54	27
841	Dipterocarpaceae	*Shorea superba* Foxw.	10	35.7	13.14	12	0	0	0	0	0	0	1	16	57	13
1060	Meliaceae	*Aglaia meliosmoides* Craib	10	35.7	25.74	9	19	0	2	0	3	0	3	8	38	17
957	Polygalaceae	*Xanthophyllum velutinum* Chodat	10	35.5	27.82	17	12	0	0	0	7	2	0	12	37	14
1187	Myristicaceae	*Gymnacranthera farquhariana* Hook.f. & Thomson Warb. var. *farquhariana*	10	35.5	30.54	11	6	0	1	16	3	1	16	1	37	10
1432	Orchidaceae	*Spathoglottis plicata* Blume	10	35.5	12.22	17	0	0	0	0	0	0	0	12	60	11
1145	Caesalpiniaceae	*Sindora beccariana* Baker ex De Wit	10	35.3	16.09	4	0	0	22	0	0	0	2	8	48	16

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1292	Moraceae	*Ficus pellucidopunctata* Griff.	10	35.3	24.64	7	0	0	0	2	11	1	2	20	40	16
1062	Meliaceae	*Aglaia odoratissima* Blume	10	35.2	25.98	16	7	0	0	3	0	0	5	9	40	20
988	Mimosaceae	*Archidendron ellipticum* Blume I.C.Nielsen	10	35	26.29	10	17	0	0	0	5	0	4	8	38	17
578	Fagaceae	*Lithocarpus urceolaris* Jack Merr.	10	34.9	20.52	7	0	0	0	0	0	0	6	12	45	29
773	Dipterocarpaceae	*Shorea argentifolia* Symington	10	34.9	19.42	13	32	0	0	0	2	0	2	0	43	8
1099	Meliaceae	*Sandoricum koetjape* Burm.f. Merr.	10	34.9	23.28	14	3	0	0	0	2	2	4	20	41	14
1165	Loranthaceae	*Helixanthera coccinea* Jack Danser	10	34.9	20.47	7	0	0	1	14	2	0	6	15	44	11
913	Ixonanthaceae	*Ixonanthes reticulata* Jack	10	34.7	28.47	12	4	0	1	0	8	3	6	25	36	5
584	Passifloraceae	*Adenia cordifolia* Blume Engl.	10	34.6	23.34	8	1	0	0	18	0	0	1	10	41	20
1085	Meliaceae	*Dysoxylum brachybotrys* Merr.	10	34.5	18.57	0	25	0	0	2	9	1	15	3	44	2
163	Euphorbiaceae	*Mallotus miquelianus* (Scheff.) Boerl.	10	34.3	25.95	8	2	0	0	2	6	0	4	26	38	15
1230	Myristicaceae	*Myristica guatteriifolia* A.DC.	10	34.3	9.44	4	0	0	0	0	0	0	0	7	67	23
524	Icacinaceae	*Sarcostigma paniculata* Pierre	10	34.2	26.51	11	0	0	0	1	11	2	6	27	38	5
747	Dipterocarpaceae	*Hopea cernua* Teijsm. & Binn.	10	34.1	20.3	14	2	0	0	0	0	0	5	4	46	29
1052	Meliaceae	*Aglaia grandis* Korth. ex Miq.	10	34.1	14.32	29	0	0	0	0	0	0	1	2	56	12
1308	Moraceae	*Ficus treubii* King	10	34.1	23.46	26	23	0	0	0	0	0	0	0	40	10
1295	Moraceae	*Ficus recurva* Blume var. *pedicellata* Corner	10	34	19.04	15	33	0	0	1	0	1	0	3	42	6
86	Euphorbiaceae	*Aporosa elmeri* Merr.	10	33.9	25.46	17	12	1	0	0	0	0	11	4	39	15
1043	Sapindaceae	*Trigonachras acuta* Radlk.	10	33.9	10.63	7	0	0	0	0	0	1	11	1	64	16
553	Fagaceae	*Castanopsis motleyana* King	10	33.8	15.9	9	3	0	2	8	0	1	0	12	49	16
732	Dipterocarpaceae	*Dipterocarpus palembanicus* Slooten	10	33.8	10.46	0	16	0	0	1	1	5	0	5	59	14
51	Annonaceae	*Polyalthia tenuipes* Merr.	10	33.7	7.08	11	0	0	0	0	0	0	0	7	77	5
130	Euphorbiaceae	*Macaranga bancana* Miq. Müll.Arg.	10	33.7	21.03	9	6	1	0	1	7	5	3	19	42	8
358	Flacourtiaceae	*Homalium foetidum* Roxb. Benth.	10	33.7	24.97	8	7	0	0	0	8	1	7	19	38	13
170	Euphorbiaceae	*Melanolepis multiglandulosa* Reinw. ex Blume Rchb.f. & Zoll.	10	33.4	15.36	15	0	0	0	0	0	1	0	5	52	27
348	Erythroxylaceae	*Erythroxylum ecarinatum* Burck	10	33.4	8.64	1	6	0	0	3	0	0	2	1	68	20
772	Dipterocarpaceae	*Shorea andulensis* P.S.Ashton	10	33.4	12.74	5	37	0	0	0	3	2	0	0	50	3
1101	Meliaceae	*Walsura aherniana* Perk.	10	33.3	8.5	23	0	0	0	0	0	0	0	4	71	2
380	Flacourtiaceae	*Scolopia spinosa* Roxb. Warb.	10	33.1	23.57	7	0	0	0	0	0	0	16	12	41	24
421	Celastraceae	*Loeseneriella pauciflora* DC. A.C.Sm.	10	33.1	12.01	0	0	0	3	10	0	0	0	19	55	13
905	Menispermaceae	*Hypserpa nitida* Miers	10	33.1	11.06	9	0	0	0	0	0	0	0	14	60	16
217	Dilleniaceae	*Dillenia excelsa* Jack Martelli var. *pubescens* Corner Masam.	10	33	11.18	5	0	0	4	0	0	0	0	13	59	19
426	Celastraceae	*Lophopetalum rigidum* Ridl.	10	33	21.31	8	22	0	1	22	4	0	1	1	39	0
968	Chrysobalaceae	*Licania splendens* Korth. Prance	10	33	23.1	5	1	0	0	4	0	0	7	26	40	17
992	Mimosaceae	*Entada rheedii* Spreng.	10	33	22.15	13	0	0	0	0	0	6	0	15	42	24
1056	Meliaceae	*Aglaia leptantha* Miq.	10	33	24.15	23	0	0	2	4	0	0	6	9	41	15
745	Dipterocarpaceae	*Hopea beccariana* Burck	10	32.9	22.77	11	6	0	0	19	0	0	3	5	40	15
346	Erythroxylaceae	*Erythroxylum burmanicum* Griff.	10	32.7	17.05	4	5	0	0	0	0	1	0	19	46	24
641	Symplocaceae	*Symplocos celastrifolia* Griff. ex C.B.Clarke	10	32.7	21.58	5	3	0	4	17	0	0	0	23	40	9
496	Loganiaceae	*Fagraea fragrans* Roxb.	10	32.6	29.98	3	7	0	0	18	6	1	5	12	34	14
844	Dipterocarpaceae	*Shorea venulosa* G.H.S.Wood ex Meijer	10	32.6	11.32	11	17	0	0	0	0	0	0	8	56	8
334	Connaraceae	*Connarus odoratus* Hook.f.	10	32.5	22.39	3	3	0	0	5	23	1	4	9	40	13
937	Polygalaceae	*Xanthophyllum adenotus* Miq. var. *arsatii* C.E.C.Fisch. W.J.de Wilde & Duyfjes	10	32.4	22.49	1	19	0	1	12	10	0	11	5	38	3
56	Annonaceae	*Popowia tomentosa* Maingay ex Hook.f. & Thomson	10	32.3	13.55	6	0	0	0	0	0	0	24	0	56	15
819	Dipterocarpaceae	*Shorea ovalis* Korth. Blume subsp. *sarawakensis* P.S.Ashton	10	32.3	12.69	0	0	0	0	0	2	0	34	0	58	6
574	Fagaceae	*Lithocarpus nieuwenhuisii* Seem. A.Camus	10	32.2	20.52	19	19	0	0	0	0	0	0	5	42	15
692	Anacardiaceae	*Melanochyla elmeri* Merr.	10	32.1	15.72	7	0	0	0	3	0	0	1	5	50	35
263	Thymelaeaceae	*Aquilaria beccariana* Tiegh.	10	32	22.99	17	8	0	0	2	0	5	5	2	40	21
305	Burseraceae	*Dacryodes costata* A.W.Benn. H.J.Lam	10	32	12.95	14	0	0	0	5	0	0	0	1	55	25
716	Dipterocarpaceae	*Dipterocarpus caudiferus* Merr.	10	32	26.56	5	6	0	9	2	0	0	4	17	36	22
1385	Orchidaceae	*Cymbidium finlaysonianum* Lindl.	10	32	12.66	1	0	0	0	20	1	0	21	3	55	0
309	Burseraceae	*Dacryodes macrocarpa* King H.J.Lam var. *macrocarpa*	10	31.9	12.78	7	0	0	0	7	0	3	0	4	55	24
757	Dipterocarpaceae	*Hopea nervosa* King	10	31.9	13.94	0	37	0	0	0	10	1	1	3	45	2

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1109	Caesalpiniaceae	*Bauhinia finlaysoniana* Benth. Graham ex Baker var. *leptopus* Perkins K.Larsen & S.S.Larsen	10	31.9	14.64	32	0	0	0	0	0	0	14	0	54	0
15	Annonaceae	*Goniothalamus woodii* Salleh	10	31.7	10.03	6	0	0	0	0	0	0	1	22	61	11
508	Simaroubaceae	*Irvingia malayana* Oliv.	10	31.6	9.54	0	0	0	0	11	0	0	0	8	62	19
1263	Moraceae	*Ficus delosyce* Corner	10	31.6	8.07	0	0	0	0	0	0	1	17	0	71	10
173	Euphorbiaceae	*Neoscortechinia angustifolia* Airy Shaw Welzen	10	31.5	12.02	7	0	0	0	0	0	0	0	21	55	16
1042	Sapindaceae	*Pometia pinnata* J.R.Forst. & G.Forst. f. *alnifolia* Blume M.Jacobs	10	31.5	21.17	17	0	0	0	0	2	4	0	19	41	16
1296	Moraceae	*Ficus recurva* Blume var. *ribesioides* Wall. King	10	31.5	20.33	12	0	0	0	0	4	0	22	6	44	12
74	Euphorbiaceae	*Antidesma neurocarpum* Miq. var. *hosei* Pax & K.Hoffm. Petra Hoffm.	10	31.4	27	25	13	5	0	1	3	0	9	3	36	5
219	Dilleniaceae	*Dillenia luzoniensis* Martelli	10	31.4	9.43	8	0	0	9	1	0	0	4	3	63	13
1041	Sapindaceae	*Paranephelium xestophyllum* Miq.	10	31.4	28.82	14	0	0	0	0	3	0	6	16	36	25
1201	Myristicaceae	*Horsfieldia ridleyana* King Warb.	10	31.2	13.01	0	20	0	4	8	6	7	0	7	48	1
1212	Myristicaceae	*Knema laurina* Blume Warb.	10	31.2	17.21	2	0	0	0	6	0	0	8	5	47	32
107	Euphorbiaceae	*Baccaurea polyneura* Hook.f.	10	31.1	26.83	14	2	0	0	6	0	0	10	5	38	25
284	Xyridaceae	*Xyris complanata* R.Br.	10	31.1	6.53	0	0	0	1	0	0	1	1	10	76	11
314	Burseraceae	*Dacryodes rostrata* Blume H.J.Lam f. *rostrata*	10	31	29.74	4	2	0	7	10	0	0	12	13	34	17
1304	Moraceae	*Ficus sumatrana* Miq. Miq.	10	31	29.79	30	3	3	0	7	0	0	2	0	36	19
706	Labiatae	*Gomphostemma microcalyx* Prain.	10	30.9	9.86	14	0	0	0	0	24	0	0	0	61	0
1170	Loranthaceae	*Lepeostegeres beccarii* King Gamble	10	30.9	24.03	22	3	0	0	1	0	0	0	5	40	28
115	Euphorbiaceae	*Cephalomappa beccariana* Baill. var. *hosei* Airy Shaw	10	30.7	17.34	0	1	0	0	14	1	2	21	5	46	10
433	Celastraceae	*Salacia chinensis* L.	10	30.7	25.53	2	5	0	19	8	1	1	6	13	35	11
541	Violaceae	*Rinorea horneri* Korth. Kuntze	10	30.7	18.93	21	23	0	0	0	0	0	0	5	42	9
250	Gnetaceae	*Gnetum latifolium* Blume var. *latifolium*	10	30.6	22.01	2	2	0	17	14	0	0	3	17	37	7
1114	Caesalpiniaceae	*Bauhinia kockiana* Korth. var. *calcicola* K.Larsen & S.S.Larsen	10	30.6	18.37	3	0	0	0	0	4	0	37	3	46	8
1218	Myristicaceae	*Knema oblongata* Merr. subsp. *oblongata* Merr.	10	30.6	22.54	15	0	0	0	0	0	2	1	13	40	29
376	Flacourtiaceae	*Ryparosa hullettii* King	10	30.5	22.78	16	0	0	0	2	5	0	7	16	40	14
1090	Meliaceae	*Dysoxylum pachyrhache* Merr.	10	30.5	22.13	16	0	0	0	0	3	1	1	17	40	21
1227	Myristicaceae	*Myristica cinnamomea* King	10	30.5	29.38	13	26	0	0	1	6	2	1	8	32	11
846	Dipterocarpaceae	*Vatica albiramis* Slooten	10	30.4	11.94	26	0	0	0	0	0	0	0	1	57	16
908	Monimiaceae	*Kibara coriacea* Blume Tul.	10	30.4	26.7	21	4	0	0	3	0	5	0	8	37	23
1140	Caesalpiniaceae	*Intsia palembanica* Miq.	10	30.4	14.92	2	0	0	0	4	0	0	2	27	48	17
58	Annonaceae	*Pseuduvaria pamattonis* Miq. Y.C.F.Su & R.M.K.Saunders	10	30.3	32.35	17	0	1	0	0	10	2	19	1	34	16
258	Sonneratiaceae	*Sonneratia alba* Sm.	10	30.3	16.71	0	1	0	10	11	4	0	1	28	42	2
391	Proteaceae	*Helicia petiolaris* Benn.	10	30.3	26.45	26	10	0	0	12	0	0	6	3	36	7
269	Thymelaeaceae	*Gonystylus keithii* Airy Shaw	10	30.2	19.9	10	6	0	0	0	0	1	0	13	42	29
372	Flacourtiaceae	*Osmelia philippina* Turcz. Benth.	10	30.2	26.75	8	17	0	6	6	8	4	3	7	33	7
632	Crypteroniaceae	*Crypteronia griffithii* C.B.Clarke	10	30.2	24.66	9	9	0	5	0	0	3	1	3	37	34
1147	Caesalpiniaceae	*Sindora irpicina* de Wit	10	30.1	20.26	0	1	0	0	0	1	5	7	10	41	34
1046	Meliaceae	*Aglaia argentea* Blume	10	30	28.11	6	0	0	0	18	3	0	13	12	35	12
1271	Moraceae	*Ficus fistulosa* Reinw.	10	30	31.65	9	18	1	0	11	4	0	11	6	32	8
221	Dilleniaceae	*Dillenia suffruticosa* Griff. Martelli	10	29.9	28.74	6	20	0	0	9	2	0	2	21	32	8
1027	Sapindaceae	*Lepisanthes fruticosa* Roxb. Leenh.	10	29.9	29.92	19	3	2	3	3	2	3	9	8	34	14
1093	Meliaceae	*Lansium domesticum* Jack	10	29.9	31.41	10	0	0	0	0	6	1	15	18	33	17
9	Annonaceae	*Disepalum anomalum* Hook.f.	10	29.8	19.58	4	6	0	1	5	2	20	5	7	40	10
363	Flacourtiaceae	*Hydnocarpus borneensis* Sleumer	10	29.8	25.25	24	0	0	0	0	3	1	1	7	38	26
660	Symplocaceae	*Symplocos polyandra* Blume Brand	10	29.8	4.22	0	0	0	0	0	0	0	0	6	94	0
713	Dipterocarpaceae	*Dipterocarpus applanatus* Sloot.	10	29.8	13.17	0	0	0	0	0	7	0	4	29	50	9
728	Dipterocarpaceae	*Dipterocarpus kunstleri* King	10	29.8	21.98	10	4	0	0	0	0	1	1	8	40	36
771	Dipterocarpaceae	*Shorea almon* Foxw.	10	29.7	6.41	0	0	0	0	0	0	0	0	14	74	12
810	Dipterocarpaceae	*Shorea maxwelliana* King	10	29.7	27.8	5	24	0	0	1	1	2	4	10	33	20
1017	Sapindaceae	*Dimocarpus fumatus* Blume Leenh. subsp. *fumatus*	10	29.7	25.14	18	0	0	0	1	0	0	3	14	38	26
1037	Sapindaceae	*Nephelium meduseum* Leenh.	10	29.7	24.19	1	30	0	1	6	8	7	0	12	33	3
1092	Meliaceae	*Heynea trijuga* Roxb. in Sims	10	29.7	20.18	24	0	0	0	0	0	0	5	7	43	21

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
939	Polygalaceae	*Xanthophyllum ellipticum* Korth. ex Miq.	10	29.6	28.75	7	9	0	1	24	3	0	3	13	32	7
1258	Moraceae	*Ficus binnendykii* Miq. Miq.	10	29.6	10.06	0	0	0	0	0	0	0	6	26	58	10
494	Loganiaceae	*Fagraea borneensis* Scheff.	10	29.5	24.2	25	4	0	0	7	0	1	0	2	38	24
808	Dipterocarpaceae	*Shorea macroptera* Dyer subsp. *sandakanensis* Symington P.S.Ashton	10	29.5	18.32	2	0	0	0	7	1	1	4	16	43	28
39	Annonaceae	*Polyalthia cauliflora* Hook.f. & Thomson var. *beccarii* Hook.f. & Thomson J.Sinclair	10	29.3	18.02	6	35	0	0	4	5	0	8	1	39	0
77	Euphorbiaceae	*Antidesma polystylum* Airy Shaw	10	29.3	6.91	0	4	0	0	0	4	0	5	10	71	7
279	Trigoniaceae	*Trigoniastrum hypoleucum* Miq.	10	29.3	32.05	16	5	0	7	7	1	1	9	10	32	13
566	Fagaceae	*Lithocarpus gracilis* Korth. Soepadmo	10	29.3	26.52	8	20	0	0	4	0	0	3	10	34	20
615	Taccaceae	*Tacca integrifolia* Ker Gawl.	10	29.3	25.17	9	15	1	0	0	18	0	2	11	35	9
238	Dilleniaceae	*Tetracera scandens* L. Merr.	10	29.2	34.21	2	4	0	0	0	5	2	21	21	31	13
525	Icacinaceae	*Stemonurus grandifolius* Becc.	10	29.2	26.85	17	10	0	0	8	0	2	0	8	35	21
625	Bignoniaceae	*Dolichandra spathacea* L.f. K.Schum.	10	29.2	7.74	0	0	0	0	1	0	0	1	21	66	11
932	Podocarpaceae	*Podocarpus polystachyus* R.Br.	10	29.1	13.65	0	0	0	4	5	1	1	13	25	49	2
248	Gnetaceae	*Gnetum gnemonoides* Brongn.	10	29	28.83	9	24	0	0	9	6	2	0	5	31	14
820	Dipterocarpaceae	*Shorea ovata* Dyer ex Brandis	10	29	9.64	0	7	0	0	18	2	0	0	16	56	0
1032	Sapindaceae	*Mischocarpus pentapetalus* Roxb. Radlk.	10	29	28.6	22	0	0	0	0	0	1	2	14	35	25
254	Gnetaceae	*Gnetum neglectum* Blume	10	28.9	30.79	11	0	0	10	0	0	1	10	11	33	24
321	Burseraceae	*Santiria laevigata* Blume f. laevigata	10	28.9	31.36	13	7	0	0	5	11	3	3	19	31	8
621	Hypericaceae	*Cratoxylum formosum* Jack Dyer	10	28.9	29.25	3	13	0	0	0	9	7	8	11	32	16
1231	Myristicaceae	*Myristica lowiana* King	10	28.7	14.7	0	0	0	0	44	2	0	2	4	45	2
35	Annonaceae	*Orophea sarawakensis* P.J.A.Kessler	10	28.6	8.85	1	6	0	0	0	0	0	0	14	61	18
385	Malpighiaceae	*Tristellateia australasiae* A.Rich.	10	28.6	7.96	0	0	0	0	9	0	0	17	4	67	3
495	Loganiaceae	*Fagraea elliptica* Roxb.	10	28.6	21.18	19	7	2	0	20	5	2	2	1	39	3
1189	Myristicaceae	*Horsfieldia brachiata* King Warb.	10	28.5	30.43	28	8	1	0	0	2	0	6	3	34	18
134	Euphorbiaceae	*Macaranga conifera* Zoll. Müll.Arg.	10	28.4	31.39	5	0	0	0	6	1	1	12	15	32	29
320	Burseraceae	*Santiria laevigata* Blume f. *glabrifolia* Engl. H.J.Lam	10	28.4	32.2	9	6	0	4	16	2	4	1	17	30	12
545	Fagaceae	*Castanopsis borneensis* King	10	28.4	8.79	2	0	0	20	0	0	0	2	15	58	1
870	Araliaceae	*Arthrophyllum diversifolium* Blume	10	28.4	34.78	13	6	2	0	20	5	1	6	8	30	9
610	Cyperaceae	*Scleria sumatrensis* Retz.	10	28.3	13.66	0	0	0	3	0	2	0	0	36	47	12
20	Annonaceae	*Meiogyne virgata* Blume Miq.	10	28.2	29.91	17	0	0	0	1	0	0	5	18	33	26
323	Burseraceae	*Santiria mollis* Engl.	10	28.2	33.04	8	5	0	0	0	6	13	1	18	30	19
501	Loganiaceae	*Strychnos ignatii* P.J.Bergius	10	28.1	35.27	6	7	0	0	0	14	1	14	14	30	13
673	Anacardiaceae	*Buchanania arborescens* Blume Blume	10	28.1	27.41	8	0	0	0	17	0	0	5	17	34	20
904	Menispermaceae	*Fibraurea tinctoria* Lour.	10	28.1	14.97	7	0	0	0	30	0	0	4	6	46	8
1035	Sapindaceae	*Nephelium cuspidatum* Blume var. *robustum* Radlk. Leenh.	10	28.1	10.87	17	0	0	0	0	0	0	2	0	57	23
1063	Meliaceae	*Aglaia oligophylla* Miq.	10	28	14.14	17	0	0	0	0	0	0	1	4	50	28
1216	Myristicaceae	*Knema membranifolia* H.J.P.Winkl.	10	28	21.92	3	14	0	0	0	14	5	5	2	37	20
241	Flagellariaceae	*Flagellaria indica* L.	10	27.8	27.26	2	2	0	13	15	0	0	5	24	32	7
620	Hypericaceae	*Cratoxylum cochinchinense* Lour. Blume	10	27.8	20.64	8	0	0	0	1	6	2	6	24	39	14
1280	Moraceae	*Ficus inaequipetiolata* Merr.	10	27.8	9.93	13	0	0	0	0	0	0	1	20	58	9
229	Dilleniaceae	*Saurauia horrida* Hook.f.	10	27.7	21.56	26	29	1	0	0	2	0	0	0	37	5
408	Rhizophoraceae	*Rhizophora apiculata* Blume	10	27.7	8.03	0	0	0	1	0	0	0	0	28	62	9
278	Thymelaeaceae	*Wikstroemia tenuiramis* Miq.	10	27.6	22.42	20	11	4	0	0	1	0	5	11	37	10
353	Flacourtiaceae	*Casearia rugulosa* Blume	10	27.6	27.41	20	10	0	0	0	0	0	12	8	34	17
377	Flacourtiaceae	*Ryparosa kostermansii* Sleumer	10	27.6	20.83	5	0	0	0	0	8	0	32	4	41	11
663	Ulmaceae	*Gironniera hirta* Ridl.	10	27.6	17.63	3	21	0	0	4	1	4	0	8	40	20
166	Euphorbiaceae	*Mallotus peltatus* (Geisel.) Müll.Arg.	10	27.5	33.74	7	11	0	11	8	3	0	5	10	29	16
661	Symplocaceae	*Symplocos tricoccata* Noot.	10	27.5	15.04	37	0	0	0	0	0	0	0	6	48	9
307	Burseraceae	*Dacryodes laxa* A.W.Benn. H.J.Lam	10	27.4	22.11	11	16	0	4	16	0	0	0	4	36	13
866	Dipterocarpaceae	*Vatica vinosa* P.S.Ashton	10	27.4	12.32	8	6	0	0	0	0	0	8	0	52	26
944	Polygalaceae	*Xanthophyllum macrophyllum* Baker	10	27.3	20.6	13	29	2	0	0	12	0	4	0	36	3
1262	Moraceae	*Ficus crassiramea* Miq.	10	27.3	26.24	0	4	0	9	2	3	11	7	19	32	14
897	Opiliaceae	*Champereia manillana* Blume Merr.	10	27.2	11.94	4	0	0	0	4	3	0	12	12	52	13
1025	Sapindaceae	*Harpullia arborea* Radlk.	10	27.2	20.42	19	0	0	0	0	2	0	1	8	40	29
1067	Meliaceae	*Aglaia silvestris* M.Roem. Merr.	10	27.2	34.6	12	7	0	0	1	1	0	6	23	29	21

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1117	Caesalpiniaceae	*Bauhinia lambiana* Baker	10	27.2	6.21	0	1	0	0	0	0	2	0	2	73	22
1207	Myristicaceae	*Knema korthalsii* Warb. subsp. *rimosa* W.J.de Wilde	10	27.2	7.05	12	0	0	0	0	0	0	0	15	68	4
81	Euphorbiaceae	*Antidesma venenosum* J.J.Sm.	10	27.1	36.69	22	12	6	0	7	5	0	8	1	29	11
534	Oxalidaceae	*Sarcotheca diversifolia* Miq. Hallier f.	10	27.1	35.37	10	7	0	0	0	10	3	5	10	29	26
616	Taccaceae	*Tacca palmata* Blume	10	27.1	18.42	4	0	1	23	0	4	3	7	15	39	5
1001	Rosaceae	*Prunus grisea* C.Muell. Kalkman var. *tomentosa* Koord. & Valeton Kalkman	10	27.1	13.5	10	31	0	0	0	1	0	10	0	46	2
1094	Meliaceae	*Pseudoclausena chrysogyne* Miq. T.Clark	10	27.1	23.13	7	0	0	3	7	0	0	4	21	36	22
7	Annonaceae	*Cyathostemma excelsum* Hook.f. & Thomson J.Sinclair	10	27	27.05	9	12	0	2	0	0	1	6	8	33	27
236	Dilleniaceae	*Tetracera akara* Burm.f. Merr.	10	27	29.7	19	8	0	0	0	0	1	7	2	33	31
342	Dichapetalaceae	*Dichapetalum gelonioides* Roxb. Engl. subsp. *pilosum* Leenh.	10	27	16.92	23	0	0	4	0	0	0	0	1	45	27
479	Ericaceae	*Vaccinium bancanum* Miq. var. *bancanum*	10	27	36.5	22	19	3	9	2	3	1	0	6	27	7
1023	Sapindaceae	*Guioa pterorhachis* Welzen	10	27	5.47	6	0	0	0	0	0	8	4	3	78	2
1232	Myristicaceae	*Myristica papyracea* J.Sinclair	10	27	23.48	0	0	0	0	0	4	3	6	15	36	35
549	Fagaceae	*Castanopsis foxworthyi* Schottky	10	26.9	20.54	23	12	0	0	13	0	1	0	8	38	5
687	Anacardiaceae	*Mangifera decandra* Ding Hou	10	26.9	14.27	0	3	0	0	0	0	18	0	10	45	24
180	Euphorbiaceae	*Pimelodendron zoanthogyne* J.J.Sm.	10	26.8	18.57	1	36	0	0	0	5	0	5	0	37	14
1250	Moraceae	*Artocarpus rigidus* Blume subsp. *rigidus*	10	26.8	19.4	11	0	0	3	8	0	0	2	12	40	24
681	Anacardiaceae	*Gluta oba* Merr. Ding Hou	10	26.7	9.86	0	1	0	3	4	7	1	3	21	54	5
205	Convolvulaceae	*Erycibe borneensis* Merr. Hoogland var. *borneensis*	10	26.6	6.37	9	0	0	0	0	0	0	2	7	72	10
316	Burseraceae	*Dacryodes rugosa* Blume H.J.Lam var. *rugosa*	10	26.6	28.38	10	30	0	0	0	2	1	2	7	30	18
742	Dipterocarpaceae	*Dryobalanops lanceolata* Burck	10	26.6	27.19	16	11	0	0	0	0	2	0	8	33	29
847	Dipterocarpaceae	*Vatica badiifolia* P.S.Ashton	10	26.6	4.96	0	8	0	0	1	0	2	4	0	80	4
1330	Orchidaceae	*Bulbophyllum acuminatum* Ridl.	10	26.6	9.89	9	2	4	0	0	0	12	6	0	56	11
749	Dipterocarpaceae	*Hopea dryobalanoides* Miq.	10	26.5	21.18	5	21	0	0	0	1	0	3	3	36	30
1193	Myristicaceae	*Horsfieldia fragillima* Airy Shaw	10	26.5	23.39	0	13	0	18	6	9	2	12	4	33	3
113	Euphorbiaceae	*Bridelia insulana* Hance	10	26.3	29.29	10	0	0	2	2	4	1	6	24	32	18
207	Convolvulaceae	*Erycibe grandifolia* Merr.	10	26.3	12.95	1	0	0	38	0	0	0	6	1	45	9
577	Fagaceae	*Lithocarpus sericobalanos* E.F.Warb.	10	26.3	19.75	11	10	0	0	2	12	2	1	1	38	23
1082	Meliaceae	*Chisocheton sarawakanus* DC. Harms	10	26.3	38.28	10	8	0	3	1	13	6	6	9	27	17
940	Polygalaceae	*Xanthophyllum flavescens* Roxb.	10	26.2	37.62	17	22	0	0	3	2	0	8	9	27	12
359	Flacourtiaceae	*Homalium grandiflorum* Benth.	10	26.1	18.57	14	0	1	0	0	24	2	6	2	40	11
554	Fagaceae	*Castanopsis oviformis* Soepadmo	10	26.1	27.03	16	9	0	0	5	0	3	2	19	32	14
694	Anacardiaceae	*Parishia maingayi* Hook.f.	10	26.1	33.97	6	16	0	10	2	4	3	2	16	27	15
855	Dipterocarpaceae	*Vatica micrantha* Slooten	10	26.1	33.65	6	26	2	3	10	9	5	2	1	27	9
1195	Myristicaceae	*Horsfieldia irya* Gaertn. Warb.	10	26.1	29.33	2	0	0	5	13	1	2	5	21	30	21
1279	Moraceae	*Ficus heteropleura* Blume	10	26.1	34.03	14	8	1	0	5	1	4	10	9	29	19
247	Gnetaceae	*Gnetum gnemon* L. var. *tenerum* Markgr.	10	26	21.02	0	16	0	4	7	4	16	2	6	34	11
941	Polygalaceae	*Xanthophyllum havilandii* Chodat	10	26	4.09	0	0	0	0	0	0	1	1	0	90	8
1050	Meliaceae	*Aglaia elliptica* Blume	10	26	36.95	15	25	0	0	1	2	0	8	8	27	14
1194	Myristicaceae	*Horsfieldia grandis* Hook.f. Warb.	10	26	29.89	7	23	0	0	11	7	3	3	9	29	7
1303	Moraceae	*Ficus spathulifolia* Corner	10	26	21.81	0	8	0	9	30	3	0	1	14	33	1
72	Euphorbiaceae	*Antidesma montanum* Blume var. *montanum*	10	25.9	26.2	5	6	0	7	10	2	2	3	16	32	18
177	Euphorbiaceae	*Omphalea bracteata* (Blanco) Merr.	10	25.9	9.27	5	0	0	0	0	0	0	6	11	58	21
422	Celastraceae	*Lophopetalum beccarianum* Pierre	10	25.9	28.7	14	25	0	0	5	0	2	1	4	30	19
671	Anacardiaceae	*Androtium astylum* Stapf	10	25.9	23.87	2	8	0	10	19	4	1	13	9	33	1
1126	Caesalpiniaceae	*Caesalpinia oppositifolia* Hattink	10	25.9	17.51	4	0	0	0	0	1	0	33	7	43	11
1309	Moraceae	*Ficus uniglandulosa* Wall.	10	25.9	35.13	14	22	4	0	13	7	0	2	3	27	9
738	Dipterocarpaceae	*Dryobalanops aromatica* Gaertn.	10	25.8	4.1	0	0	0	0	9	0	0	0	0	88	3
13	Annonaceae	*Goniothalamus tapis* Miq.	10	25.7	16.98	1	31	0	4	11	7	0	1	4	37	5
388	Proteaceae	*Helicia excelsa* Roxb. Blume	10	25.7	32.96	23	6	1	0	0	3	1	3	17	30	15
398	Rhizophoraceae	*Bruguiera sexangula* Lour. Poir.	10	25.7	11.31	0	0	0	0	11	0	0	1	34	49	6
715	Dipterocarpaceae	*Dipterocarpus caudatus* Foxw. subsp. *penangianus* Foxw. P.S.Ashton	10	25.7	13.98	0	25	0	0	7	0	8	4	3	42	10
80	Euphorbiaceae	*Antidesma tomentosum* Blume	10	25.6	36.83	15	20	1	0	2	4	4	7	9	27	13

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
603	Cyperaceae	*Paramapania radians* C.B.Clarke Uittien	10	25.6	23.25	4	25	0	0	0	27	2	2	3	32	5
858	Dipterocarpaceae	*Vatica oblongifolia* Hook.f. subsp. *multinervosa*	10	25.6	12.85	0	15	0	0	0	5	4	0	4	46	25
875	Araliaceae	*Schefflera elliptica* Blume Harms	10	25.6	23.62	26	0	3	5	12	0	0	1	8	35	9
1053	Meliaceae	*Aglaia korthalsii* Miq.	10	25.6	21.29	26	0	0	0	0	0	0	4	2	39	29
1143	Caesalpiniaceae	*Peltophorum racemosum* Merr.	10	25.6	13.51	0	0	0	0	0	2	0	32	2	49	14
535	Oxalidaceae	*Sarcotheca glauca* Hallier	10	25.5	20.64	1	1	0	10	30	0	0	0	22	34	2
539	Violaceae	*Rinorea anguifera* Lour. Kuntze	10	25.5	33.29	6	20	0	6	8	16	2	9	4	27	2
1200	Myristicaceae	*Horsfieldia polyspherula* Hook.f. emend. King J.Sinclair var. *sumatrana* Miq. W.J.de Wilde	10	25.5	34.64	12	22	0	0	0	1	2	8	16	27	12
1247	Moraceae	*Artocarpus nitidus* Trécul	10	25.5	32.44	14	4	0	0	3	2	3	5	21	29	18
1315	Moraceae	*Prainea limpato* Miq. Beumée ex K.Heyne subsp. *limpato*	10	25.5	20.61	2	3	0	0	0	0	11	2	18	36	28
64	Annonaceae	*Woodiellantha* sympetala Merr. S.Rauschert	10	25.4	11.1	18	0	0	0	0	0	0	2	0	54	25
697	Anacardiaceae	*Pentaspadon motleyi* Hook.f.	10	25.4	34.43	3	19	0	5	6	3	3	8	11	27	16
1051	Meliaceae	*Aglaia forbesii* King	10	25.4	21.71	11	6	1	1	0	16	4	3	11	35	10
1255	Moraceae	*Ficus annulata* Blume	10	25.4	41.75	8	12	0	1	0	3	1	14	18	26	17
1078	Meliaceae	*Chisocheton pentandrus* Blanco Merr. subsp. *medius* Mabb.	10	25.3	22.58	22	0	0	0	0	0	0	6	9	37	26
1123	Caesalpiniaceae	*Caesalpinia crista* L.	10	25.3	11.88	0	0	0	6	15	0	0	3	21	48	8
206	Convolvulaceae	*Erycibe glomerata* Blume subsp. *angustifolia* Hallier f. Hoogland	10	25.2	8.22	18	0	0	0	0	0	1	0	0	62	18
903	Menispermaceae	*Diploclisia kunstleri* King Diels	10	25.2	16.15	18	24	0	0	0	4	0	0	7	40	6
1047	Meliaceae	*Aglaia aspera* Teijsm. & Binn.	10	25.2	13.84	0	35	0	0	0	4	3	11	6	41	0
1103	Meliaceae	*Walsura villamilii* Merr.	10	25.2	4.19	0	0	0	0	0	0	0	0	0	87	13
1259	Moraceae	*Ficus borneensis* Kochummen	10	25.2	26.73	18	3	5	0	19	14	0	3	5	32	0
257	Sonneratiaceae	*Duabanga moluccana* Blume	10	25.1	39.78	16	12	0	0	3	8	1	4	9	26	20
1281	Moraceae	*Ficus lepicarpa* Blume	10	25.1	30.5	12	22	3	0	0	10	7	2	6	28	9
88	Euphorbiaceae	*Aporosa frutescens* Blume	10	25	25.56	20	0	0	2	0	0	0	13	8	35	23
862	Dipterocarpaceae	*Vatica sarawakensis* Heim	10	25	38.59	14	18	0	5	11	1	0	5	7	26	14
1209	Myristicaceae	*Knema latericia* Elmer subsp. *albifolia* Sinclair W.J.de Wilde	10	25	24.1	4	25	0	0	0	0	2	3	7	32	26
397	Rhizophoraceae	*Bruguiera parviflora* Wight & Arn.	10	24.9	11.71	1	0	0	15	2	3	0	3	30	46	0
984	Mimosaceae	*Archidendron clypearia* Jack I.C.Nielsen	10	24.9	40.41	11	3	4	1	0	11	6	8	19	25	11
991	Mimosaceae	*Archidendron microcarpum* Benth. I.C.Nielsen	10	24.9	19.8	7	0	0	0	0	0	0	7	15	39	32
1142	Caesalpiniaceae	*Peltophorum pterocarpum* DC. Backer ex K.Heyne	10	24.9	10.28	0	0	0	3	8	0	0	3	31	51	5
190	Burmaniaceae	*Burmannia coelestis* D.Don	10	24.8	17.51	0	1	0	5	8	4	4	2	35	37	3
530	Ochnaceae	*Gomphia serrata* Gaertn. Kanis	10	24.8	35.28	15	20	0	0	11	1	1	3	8	27	13
1	Annonaceae	*Alphonsea javanica* Scheff.	10	24.7	11.81	6	0	0	0	0	0	0	0	5	50	39
290	Burseraceae	*Canarium denticulatum* Blume f. *fissistipulum* Miq. Leenh.	10	24.7	33.78	16	14	0	0	0	8	2	4	24	27	6
351	Flacourtiaceae	*Casearia grewiaefolia* Vent. var. *gelonioides*	10	24.7	16.66	29	2	0	0	0	0	1	0	2	43	22
851	Dipterocarpaceae	*Vatica havilandii* Brandis	10	24.7	15.55	0	9	0	0	11	5	0	19	13	42	2
249	Gnetaceae	*Gnetum klossii* Merr. ex Markgr.	10	24.6	28.7	6	17	1	0	0	25	4	8	9	29	2
295	Burseraceae	*Canarium littorale* Blume f. *littorale*	10	24.6	26.02	22	1	0	0	0	0	4	9	9	34	21
329	Connaraceae	*Agelaea borneensis* Hook.f. Merr.	10	24.6	37.95	9	13	0	2	6	4	5	1	12	26	23
1210	Myristicaceae	*Knema latericia* Elmer subsp. *ridleyi* Gand. W.J.de Wilde f. *olivacea* W.J.de Wilde	10	24.6	29.47	28	0	1	0	2	0	1	7	6	32	22
328	Burseraceae	*Triomma malaccensis* Hook.f.	10	24.5	16.1	5	0	0	0	0	0	0	0	14	42	39
806	Dipterocarpaceae	*Shorea leprosula* Miq.	10	24.5	41.9	6	8	0	7	0	7	2	9	19	25	19
29	Annonaceae	*Neo-uvaria acuminatissima* Miq. Airy Shaw	10	24.4	31.56	14	2	0	0	7	1	10	4	6	30	26
160	Euphorbiaceae	*Mallotus lackeyi* Elmer	10	24.4	17.02	15	0	0	0	0	0	0	1	10	42	32
612	Vitaceae	*Leea aculeata* Blume ex. Spreng.	10	24.4	36.38	15	11	0	0	0	1	1	12	9	28	22
859	Dipterocarpaceae	*Vatica oblongifolia* Hook.f. subsp. *oblongifolia*	10	24.4	23.56	3	8	0	7	0	1	18	3	11	32	19
271	Thymelaeaceae	*Gonystylus micranthus* Airy Shaw	10	24.3	16.21	0	1	0	0	28	1	0	17	9	41	3
561	Fagaceae	*Lithocarpus conocarpus* Oudem. Rehder	10	24.2	37.81	20	14	2	0	0	6	3	3	10	26	14

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
774	Dipterocarpaceae	*Shorea atrinervosa* Symington	10	24.2	39.16	9	10	0	8	11	3	1	1	17	25	15
1033	Sapindaceae	*Mischocarpus sundaicus* Blume	10	24.2	29.46	19	0	2	0	0	8	6	4	11	31	19
404	Rhizophoraceae	*Gynotroches axillaris* Blume	10	24.1	35.89	18	14	3	4	4	4	4	4	14	26	5
1122	Caesalpiniaceae	*Caesalpinia bonduc* L. Roxb.	10	24.1	7.73	0	0	0	2	2	0	0	1	19	60	16
675	Anacardiaceae	*Buchanania sessilifolia* Blume	10	24	40.84	18	14	1	1	0	1	1	9	7	26	22
1257	Moraceae	*Ficus aurita* Reinw. ex Blume	10	24	23.39	22	0	3	0	0	2	0	17	15	35	4
94	Euphorbiaceae	*Aporosa nitida* Merr.	10	23.9	38.94	8	15	0	2	1	6	5	7	8	25	22
266	Thymelaeaceae	*Gonystylus affinis* Radlk.	10	23.9	18.53	0	10	0	7	9	2	3	0	21	35	13
292	Burseraceae	*Canarium dichotomum* Blume Miq.	10	23.9	31.33	15	12	2	0	0	2	9	2	5	29	25
298	Burseraceae	*Canarium littorale* Blume f. *tomentosum* Blume Leenh.	10	23.9	30.48	13	21	0	0	5	0	2	6	6	29	18
300	Burseraceae	*Canarium merrillii* H.J.Lam	10	23.9	13.47	29	0	10	1	3	2	0	0	10	45	0
712	Dipterocarpaceae	*Dipterocarpus acutangulus* Vesque	10	23.9	40.26	12	17	0	9	1	3	3	3	11	24	16
1253	Moraceae	*Ficus adenosperma* Miq.	10	23.9	16.65	7	15	0	0	11	0	0	8	1	40	18
37	Annonaceae	*Phaeanthus splendens* Miq.	10	23.8	27.29	1	1	0	1	16	2	14	4	10	30	20
49	Annonaceae	*Polyalthia motleyana* Hook.f. Airy Shaw	10	23.8	18.74	15	22	0	0	4	0	0	6	11	36	5
1013	Sapindaceae	*Allophylus cobbe* L. Raeusch.	10	23.8	30.69	20	7	0	1	0	0	0	11	9	30	21
368	Flacourtiaceae	*Hydnocarpus sumatrana* Miq. Koord.	10	23.7	28.12	11	2	1	0	0	18	1	4	22	30	11
935	Polygalaceae	*Polygala venenosa* Juss.	10	23.7	23.98	20	4	11	0	0	29	0	4	0	33	0
945	Polygalaceae	*Xanthophyllum neglectum* Meijden	10	23.7	22.68	2	0	0	0	0	0	1	31	21	35	9
1072	Meliaceae	*Aphanamixis borneensis* Harms	10	23.7	35.04	17	18	1	1	6	0	0	11	8	27	12
1190	Myristicaceae	*Horsfieldia carnosa* Warb.	10	23.7	7.86	0	0	0	0	42	0	0	1	0	57	0
150	Euphorbiaceae	*Macaranga tanarius* L. Müll.Arg.	10	23.6	17.1	10	0	0	0	0	0	1	5	4	41	38
839	Dipterocarpaceae	*Shorea smithiana* Symington	10	23.6	12.99	1	0	0	0	2	0	0	9	8	47	33
5	Annonaceae	*Cyathocalyx magnifica* Diels	10	23.5	34.6	15	11	6	0	2	6	4	9	6	27	14
159	Euphorbiaceae	*Mallotus korthalsii* Müll.Arg.	10	23.5	32.28	8	10	0	3	0	13	8	5	11	27	15
172	Euphorbiaceae	*Moultonianthus leembruggianus* Boerl. & Koord. Steenis	10	23.5	25.37	2	24	1	0	0	5	13	3	10	29	14
365	Flacourtiaceae	*Hydnocarpus elmeri* Merr.	10	23.5	13.02	2	5	0	0	0	1	0	20	1	47	24
951	Polygalaceae	*Xanthophyllum resupinatum* Meijden	10	23.5	24.51	6	29	4	0	0	18	3	6	0	30	5
139	Euphorbiaceae	*Macaranga hypoleuca* Rchb.f. & Zoll. Müll.Arg.	10	23.4	35.23	8	0	0	5	0	5	2	10	21	27	22
322	Burseraceae	*Santiria megaphylla* Kalkman	10	23.4	12.14	0	31	0	0	9	1	6	2	4	42	5
425	Celastraceae	*Lophopetalum multinervium* Ridl.	10	23.4	30.13	12	0	0	0	24	1	3	1	16	29	13
666	Ulmaceae	*Gironniera subaequalis* Planch.	10	23.4	42.09	15	12	5	0	3	13	2	11	6	24	8
756	Dipterocarpaceae	*Hopea micrantha* Hook.f.	10	23.4	13.05	0	2	0	0	2	0	9	3	8	45	30
978	Mimosaceae	*Albizia corniculata* Lour. Druce	10	23.4	4.08	1	0	0	0	0	0	0	0	0	85	14
1256	Moraceae	*Ficus apiocarpa* Miq.	10	23.4	8.67	0	0	0	0	25	0	0	2	10	55	7
367	Flacourtiaceae	*Hydnocarpus subfalcata* Merr.	10	23.3	28.42	11	27	4	0	0	9	3	3	7	28	8
915	Linaceae	*Indorouchera griffithiana* Planch. Hallier f.	10	23.3	41.24	7	17	0	3	0	14	7	7	13	23	9
220	Dilleniaceae	*Dillenia pulchella* Jack Gilg	10	23.2	18.69	0	0	0	1	34	4	0	4	19	35	2
289	Burseraceae	*Canarium denticulatum* Blume f. *fissistipulatum* Leenh.	10	23.2	31.19	13	20	0	1	0	2	2	5	6	28	22
1272	Moraceae	*Ficus fulva* Reinw. ex Blume	10	23.2	15.94	14	13	8	0	0	22	1	1	1	39	2
105	Euphorbiaceae	*Baccaurea odoratissima* Elmer	10	23.1	34.77	23	19	2	0	0	2	5	3	6	27	14
126	Euphorbiaceae	*Homonoia riparia* Lour.	10	23.1	5.37	8	0	0	0	0	0	0	1	5	73	13
375	Flacourtiaceae	*Ryparosa hirsuta* J.J.Sm.	10	23.1	25.52	1	17	0	0	12	9	2	2	21	29	6
879	Araliaceae	*Schefflera petiolosa* Miq. Harms	10	23.1	26.89	9	19	1	0	0	12	0	18	6	30	5
87	Euphorbiaceae	*Aporosa falcifera* Hook.f.	10	23	29.07	17	0	0	0	5	0	2	10	5	31	30
296	Burseraceae	*Canarium littorale* Blume f. *pruinosum* Leenh.	10	23	30.43	12	24	3	0	0	1	5	11	0	28	17
386	Pittosporaceae	*Pittosporum ferrugineum* Aiton	10	23	10.53	20	6	0	0	0	0	0	1	12	51	10
601	Cyperaceae	*Mapania cuspidata* Miq. Uittien var. *petiolata* C.B.Clarke Uittien	10	23	20.24	8	0	0	19	0	0	7	13	1	36	17
1066	Meliaceae	*Aglaia shawiana* Merr.	10	23	5.57	9	0	0	0	0	0	0	0	0	72	19
50	Annonaceae	*Polyalthia rumphii* Blume Merr.	10	22.8	36.7	15	5	0	0	0	0	2	17	8	27	26
528	Icacinaceae	*Stemonurus umbellatus* Becc.	10	22.8	33.07	7	12	3	1	14	4	10	8	10	26	5
668	Ulmaceae	*Trema orientalis* L. Blume	10	22.8	24.64	5	1	0	13	0	0	0	9	20	32	20
61	Annonaceae	*Uvaria ovalifolia* Blume	10	22.7	19.77	16	5	1	0	6	0	12	1	3	36	20
731	Dipterocarpaceae	*Dipterocarpus pachyphyllus* Meyer	10	22.7	19.96	6	3	0	0	0	0	3	11	8	37	32
1158	Loranthaceae	*Amyema beccarii* Tiegh. Danser	10	22.7	18.72	18	0	0	0	9	0	4	11	5	38	15

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
57	Annonaceae	*Pseuduvaria borneensis* Y.C.F.Su & R.M.K.Saunders	10	22.6	27.35	17	5	0	0	0	0	4	10	6	31	26
196	Combretaceae	*Combretum nigrescens* King	10	22.6	8.29	10	0	0	0	0	0	0	0	10	58	22
200	Combretaceae	*Terminalia catappa* L.	10	22.6	12.2	0	0	0	2	16	0	0	9	24	45	5
302	Burseraceae	*Canarium odontophyllum* Miq. f. odotophyllum	10	22.6	28.27	24	3	1	0	7	0	1	10	0	31	23
812	Dipterocarpaceae	*Shorea multiflora* Burck Symington	10	22.6	39.15	4	23	1	1	8	10	5	11	7	23	8
1249	Moraceae	*Artocarpus rigidus* Blume	10	22.6	26.86	12	3	0	0	0	6	2	13	6	32	25
1273	Moraceae	*Ficus glandulifera* Wall. ex Miq. King	10	22.6	24.22	22	5	4	0	0	2	6	0	14	32	15
17	Annonaceae	*Maasia hypoleuca* Hook.f. & Thomson. Mols. P.J.A.Keßler & S.H.Rogstad	10	22.5	17.44	0	19	0	0	2	12	3	7	19	35	2
114	Euphorbiaceae	*Bridelia stipularis* L. Blume	10	22.5	7.05	15	0	0	0	0	0	0	0	9	63	13
227	Dilleniaceae	*Saurauia ferox* Korth.	10	22.4	20.89	14	29	3	0	0	17	0	2	2	32	2
679	Anacardiaceae	*Gluta beccarii* Engl. Ding Hou	10	22.4	8.97	0	12	0	0	23	0	0	1	13	50	0
975	Mimosaceae	*Acacia pseudointsia* Miq.	10	22.4	9.29	12	0	0	0	0	0	0	1	0	55	32
1054	Meliaceae	*Aglaia lawii* Wight C.J.Saldanha	10	22.4	37.6	12	16	2	0	7	10	0	2	12	24	15
602	Cyperaceae	*Mapania latifolia* Uittien	10	22.3	15.1	44	0	0	0	0	0	0	6	0	44	5
1235	Nepenthaceae	*Nepenthes gracilis* Korth.	10	22.2	34.1	6	2	0	3	19	7	1	6	22	26	10
1312	Moraceae	*Ficus xylophylla* Wall. ex Miq. Miq.	10	22.2	13.26	3	6	0	2	32	0	0	1	5	42	9
684	Anacardiaceae	*Gluta wallichii* Hook.f. Ding Hou	10	22.1	34.06	9	5	0	0	16	1	2	5	15	26	19
1058	Meliaceae	*Aglaia luzoniensis* Vidal Merr. & Rolfe	10	22.1	9.85	33	0	0	0	0	0	0	10	0	55	3
215	Dilleniaceae	*Dillenia beccariana* Martelli	10	22	25.28	1	17	0	0	1	16	1	22	8	30	4
952	Polygalaceae	*Xanthophyllum reticulatum* Chodat	10	22	11.24	18	0	0	0	0	0	1	2	0	50	29
1087	Meliaceae	*Dysoxylum cyrtobotryum* Miq.	10	22	46.95	15	4	3	0	5	14	2	8	8	23	18
1135	Caesalpiniaceae	*Cynometra ramiflora* L.	10	22	14.07	0	0	0	0	18	0	0	1	36	40	5
1196	Myristicaceae	*Horsfieldia laticostata* J.Sinclair W.J.de Wilde	10	22	31.58	0	11	0	14	12	9	6	5	14	25	3
717	Dipterocarpaceae	*Dipterocarpus confertus* Slooten	10	21.9	18.99	5	0	0	0	0	3	4	1	21	36	30
1305	Moraceae	*Ficus sundaica* Blume	10	21.9	43.57	10	4	1	0	18	6	5	10	17	23	6
129	Euphorbiaceae	*Macaranga auriculata* Merr. Airy Shaw	10	21.8	29.16	20	3	1	0	0	0	4	18	2	30	20
555	Fagaceae	*Castanopsis psilophylla* Soepadmo	10	21.8	11.55	34	0	0	0	0	0	0	0	4	49	12
1057	Meliaceae	*Aglaia leucophylla* King	10	21.8	26.64	12	25	0	0	0	2	0	17	5	30	10
1307	Moraceae	*Ficus tinctoria* G.Forst. subsp. *gibbosa* Blume Corner	10	21.8	29.56	19	26	3	0	0	12	0	5	1	27	5
34	Annonaceae	*Orophea myriantha* Merr.	10	21.7	14.48	14	0	0	0	0	0	0	18	8	44	17
203	Combretaceae	*Terminalia phellocarpa* King	10	21.7	14.19	0	27	0	0	0	8	0	9	13	38	4
252	Gnetaceae	*Gnetum leptostachyum* Blume	10	21.7	32.09	13	18	2	0	0	23	0	9	0	27	7
301	Burseraceae	*Canarium odontophyllum* Miq. f. *multifidum* H.J.Lam Leenh.	10	21.7	6.62	1	0	0	0	0	0	0	0	1	64	34
799	Dipterocarpaceae	*Shorea hypoleuca* Meijer	10	21.7	3.73	11	0	0	0	0	0	0	0	1	86	2
1181	Loranthaceae	*Scurrula ferruginea* Jack Danser	10	21.7	25.67	7	1	0	0	22	7	0	4	20	30	9
71	Euphorbiaceae	*Antidesma montanum* Blume	10	21.6	46.55	13	11	0	4	15	5	1	6	12	22	12
246	Gnetaceae	*Gnetum gnemon* L. var. *brunonianum* Griff. Markgr.	10	21.6	12.16	0	22	0	0	0	23	4	2	8	41	0
332	Connaraceae	*Connarus euphlebius* Merr. var. *euphlebius*	10	21.6	20.83	5	0	0	0	0	0	7	2	33	33	19
378	Flacourtiaceae	*Ryparosa kunstleri* King	10	21.6	16.32	0	0	0	0	0	2	1	24	25	39	10
736	Dipterocarpaceae	*Dipterocarpus validus* Blume	10	21.6	15.38	5	0	0	0	6	0	0	1	12	40	35
865	Dipterocarpaceae	*Vatica umbonata* Hook.f. Burck	10	21.6	43.79	15	11	0	2	3	2	3	5	14	23	22
6	Annonaceae	*Cyathocalyx ramuliflorus* Maingay ex Hook.f. & Thomson Scheff.	10	21.5	4.91	0	0	0	0	0	4	3	11	0	74	7
214	Datsicaceae	*Octomeles sumatrana* Miq.	10	21.5	31.39	0	11	0	0	8	3	4	9	25	26	14
1030	Sapindaceae	*Lepisanthes rubiginosa* Roxb. Leenh.	10	21.5	8.53	1	0	0	0	3	0	0	14	6	56	20
162	Euphorbiaceae	*Mallotus minimifructus* S.E.C. Sierra	10	21.4	30.28	19	21	0	0	0	2	5	8	3	28	15
700	Anacardiaceae	*Semecarpus cuneiformis* Blanco	10	21.4	16.86	14	0	0	0	6	0	0	2	1	40	38
1065	Meliaceae	*Aglaia rivularis* Merr.	10	21.3	13.58	22	0	0	0	0	0	1	0	6	44	26
1176	Loranthaceae	*Macrosolen cochinchinensis* Lour. Tiegh.	10	21.3	41.34	11	11	1	0	20	6	1	2	17	23	9
8	Annonaceae	*Cyathostemma hookeri* King	10	21.2	8.79	0	0	0	0	0	2	0	38	0	56	4
324	Burseraceae	*Santiria oblongifolia* Blume	10	21.2	34.86	5	16	0	1	6	0	6	1	18	24	24
1237	Nepenthaceae	*Nepenthes rafflesiana* Jack	10	21.2	37.89	3	7	0	16	18	5	4	2	16	23	7
503	Loganiaceae	*Strychnos villosa* A.W.Hill	10	21.1	10	1	0	0	0	0	0	0	0	20	49	29
1202	Myristicaceae	*Horsfieldia splendida* W.J.de Wilde	10	21.1	26.29	7	17	0	7	14	9	1	1	10	27	6
744	Dipterocarpaceae	*Dryobalanops rappa* Becc.	10	21	13.38	0	0	0	27	14	1	1	1	16	38	2

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
883	Liliaceae	*Peliosanthes teta* Andrews	10	21	4.01	0	0	0	0	0	2	0	0	16	79	3
1153	Boraginaceae	*Pteleocarpa lamponga* Miq. Bakh.	10	21	13.8	15	0	0	0	0	0	0	0	18	42	24
244	Gnetaceae	*Gnetum cuspidatum* Blume	10	20.9	44.82	11	21	6	1	5	10	0	11	9	21	5
864	Dipterocarpaceae	*Vatica umbonata* Hook.f. Burck subsp. *umbonata*	10	20.9	27.48	12	22	0	1	8	0	2	12	2	28	12
1220	Myristicaceae	*Knema pedicellata* W.J.de Wilde	10	20.9	10.42	0	3	0	31	3	5	0	3	10	44	1
1244	Moraceae	*Artocarpus glaucus* Blume	10	20.9	17.8	2	0	0	0	3	0	11	8	11	36	30
11	Annonaceae	*Goniothalamus roseus* Stapf	10	20.8	9.85	41	0	0	0	0	0	0	0	0	53	7
340	Connaraceae	*Rourea minor* Gaertn. Leenh.	10	20.8	47.6	11	19	0	5	3	12	2	9	8	21	11
1102	Meliaceae	*Walsura pinnata* Hassk.	10	20.8	24.66	10	1	0	0	0	4	4	16	16	31	18
175	Euphorbiaceae	*Neoscortechinia philippinensis* Merr. Welzen	10	20.7	40.03	13	17	0	0	0	4	2	9	15	23	18
198	Combretaceae	*Lumnitzera littorea* Jack Voigt	10	20.7	5.18	0	0	0	0	0	0	0	4	21	68	7
327	Burseraceae	*Santiria tomentosa* Blume	10	20.7	43.34	5	18	1	0	5	12	8	4	12	21	13
643	Symplocaceae	*Symplocos crassipes* C.B.Clarke var. *ernae* Brand Noot.	10	20.7	36.49	19	6	11	0	0	6	11	8	3	25	12
723	Dipterocarpaceae	*Dipterocarpus globosus* Vesque	10	20.7	3.24	0	6	0	0	1	0	3	2	1	88	1
802	Dipterocarpaceae	*Shorea johorensis* Foxw.	10	20.7	25.45	7	0	0	0	2	0	0	12	19	31	29
1288	Moraceae	*Ficus midotis* Corner	10	20.7	31.64	19	12	5	0	0	4	0	10	12	27	12
142	Euphorbiaceae	*Macaranga lowii* King ex Hook.f.	10	20.6	26.42	8	0	0	0	3	1	1	12	15	30	30
914	Linaceae	*Ctenolophon parvifolius* Oliv.	10	20.6	32.97	8	5	0	3	24	7	1	2	18	25	8
1018	Sapindaceae	*Dimocarpus longan* Lour. subsp. *malesianus* Leenh.	10	20.6	47.07	8	19	0	4	0	3	5	12	8	21	20
526	Icacinaceae	*Stemonurus malaccensis* Mast. Sleumer	10	20.5	48.44	15	10	1	11	3	3	5	7	10	21	13
665	Ulmaceae	*Gironniera parvifolia* Planch.	10	20.5	12.11	42	0	4	0	0	0	3	0	0	47	4
993	Mimosaceae	*Leucaena leucocephala* Lam. de Wit	10	20.5	14.17	15	0	0	0	0	0	1	0	1	42	41
1020	Sapindaceae	*Dodonaea viscosa* Jacq.	10	20.5	15.63	0	14	0	2	17	5	10	3	12	35	3
1219	Myristicaceae	*Knema pallens* W.J.de Wilde	10	20.5	7.99	1	0	0	0	0	0	0	1	1	56	41
1313	Moraceae	*Parartocarpus bracteatus* King Becc.	10	20.5	20.21	4	8	0	0	0	0	19	0	5	32	31
122	Euphorbiaceae	*Excoecaria agallocha* L.	10	20.4	4.19	0	0	0	0	0	0	0	1	20	76	3
123	Euphorbiaceae	*Gymnanthes borneensis* Pax & K.Hoffm. Esser	10	20.4	37.53	11	19	1	0	1	15	6	5	11	23	9
497	Loganiaceae	*Fagraea racemosa* Wall.	10	20.4	47.64	11	15	0	2	8	11	2	4	13	21	13
906	Menispermaceae	*Pericampylus glaucus* Lam. Merr.	10	20.4	25.34	16	0	7	0	0	13	1	3	28	29	3
46	Annonaceae	*Polyalthia igniflora* D.M.Johnson	10	20.3	7.12	18	0	0	0	0	0	0	1	0	60	20
156	Euphorbiaceae	*Mallotus caudatus* Merr.	10	20.3	14	5	0	0	0	0	5	2	2	29	40	17
294	Burseraceae	*Canarium latistipulatum* Ridl.	10	20.3	17.66	3	0	0	0	0	1	0	4	33	35	23
835	Dipterocarpaceae	*Shorea scabrida* Symington	10	20.3	23.19	0	15	0	7	23	8	3	9	5	29	2
894	Olacaceae	*Olax imbricata* Roxb.	10	20.3	39.01	1	5	0	14	0	7	4	18	11	23	16
900	Chloranthaceae	*Sarcandra glabra* Thunb. Nakai	10	20.3	30.75	23	3	18	0	0	17	0	11	0	27	0
1241	Moraceae	*Artocarpus anisophyllus* Miq.	10	20.3	42.81	10	0	0	0	23	0	0	14	13	23	16
597	Cyperaceae	*Fimbristylis littoralis* Gaudich.	10	20.2	34.24	20	8	0	0	0	1	2	8	12	26	23
887	Liliaceae	*Smilax gigantea* Merr.	10	20.2	15.44	10	0	0	7	0	0	0	13	1	40	29
240	Droseracaea	*Drosera burmanni* Vahl	10	20.1	7.6	0	0	0	0	27	0	0	2	18	53	0
401	Rhizophoraceae	*Ceriops decandra* Griff. Ding Hou	10	20.1	16.32	0	1	0	20	8	5	3	11	11	35	6
1169	Loranthaceae	*Lepeostegeres bahajensis* Korth. Miq.	10	20.1	21.55	1	18	0	0	20	2	11	2	13	29	4
1294	Moraceae	*Ficus punctata* Thunb.	10	20.1	39.08	18	12	15	0	0	8	6	8	2	23	9
100	Euphorbiaceae	*Aporosa symplocoides* Hook.f. Gage var. *chlalarocarpa* Airy Shaw Schot	10	20	13.98	37	7	0	0	0	0	0	0	1	42	13
216	Dilleniaceae	*Dillenia excelsa* Jack Martelli var. *excelsa*	10	20	42.81	2	15	0	0	12	13	4	5	21	21	8
347	Erythroxylaceae	*Erythroxylum cuneatum* Miq. Kurz	10	20	20.39	0	1	0	0	21	0	3	12	9	33	21
837	Dipterocarpaceae	*Shorea seminis* de Vriese Slooten	10	20	44.47	1	11	0	2	11	15	10	5	16	20	8
1245	Moraceae	*Artocarpus lacucha* Buch.-Ham.	10	20	48.89	7	11	0	4	1	12	6	10	15	20	13
264	Thymelaeaceae	*Aquilaria malaccensis* Lam.	10	19.9	10.61	18	0	0	0	0	0	0	0	5	48	29
110	Euphorbiaceae	*Baccaurea sumatrana* Miq. Müll.Arg.	10	19.8	49.94	14	17	0	1	1	3	3	10	14	20	16
1040	Sapindaceae	*Nephelium uncinatum* Radlk. ex Leenh.	10	19.8	37.45	7	3	0	0	3	1	11	14	15	24	22
23	Annonaceae	*Mezzettia umbellata* Becc.	10	19.7	33.19	5	10	1	7	2	15	6	10	15	24	4
188	Euphorbiaceae	*Wetria insignis* Steud. Airy Shaw	10	19.7	26.68	11	0	0	0	0	1	3	6	24	29	26
400	Rhizophoraceae	*Carallia brachiata* Lour. Merr.	10	19.7	45.05	10	14	0	0	11	6	1	5	16	21	17
1113	Caesalpiniaceae	*Bauhinia kockiana* Korth.	10	19.7	33.4	21	7	2	0	0	2	3	14	11	26	15
1154	Boraginaceae	*Tournefortia tetranda* Blume	10	19.7	15.33	21	0	0	0	0	2	0	1	4	40	33

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
16	Annonaceae	*Maasia glauca* Hassk. Mols. P.J.A.Keßler & S.H.Rogstad	10	19.6	42.64	4	20	0	0	0	7	7	7	20	21	15
563	Fagaceae	*Lithocarpus dasystachyus* Miq. Rehder	10	19.6	19.61	0	4	0	19	17	4	8	1	16	30	2
90	Euphorbiaceae	*Aporosa grandistipula* Merr.	10	19.5	47.17	6	16	1	0	5	7	10	11	9	20	16
182	Euphorbiaceae	*Sauropus bacciformis* L. Airy Shaw	10	19.5	2.87	0	0	0	0	0	0	0	1	1	93	5
779	Dipterocarpaceae	*Shorea bullata* P.S.Ashton	10	19.5	3.19	0	6	0	0	1	0	3	3	1	86	1
1144	Caesalpiniaceae	*Saraca declinata* Jack Miq.	10	19.5	45.13	9	20	1	0	0	16	2	11	5	21	13
727	Dipterocarpaceae	*Dipterocarpus kerrii* King	10	19.4	6.26	3	0	0	0	0	0	0	10	0	63	23
1119	Caesalpiniaceae	*Bauhinia semibifida* Roxb. var. *semibifida*	10	19.4	47.7	7	12	0	4	3	12	8	10	8	20	15
125	Euphorbiaceae	*Homalanthus populneus* Geiseler Pax	10	19.3	48.77	11	15	1	0	7	8	4	6	12	20	15
402	Rhizophoraceae	*Ceriops tagal* Perr. C.B.Rob.	10	19.3	8.71	0	0	0	0	1	0	0	5	26	50	18
849	Dipterocarpaceae	*Vatica dulitensis* Symington	10	19.3	24.68	27	8	2	2	3	3	4	0	5	30	16
987	Mimosaceae	*Archidendron cockburnii* I.C.Nielsen	10	19.3	9.01	12	0	0	0	0	0	0	0	0	52	36
1038	Sapindaceae	*Nephelium ramboutan-ake* Labill. Leenh.	10	19.3	37.8	22	23	2	0	0	0	1	5	8	23	16
286	Burseraceae	*Canarium asperum* Benth. var. *asperum*	10	19.2	10.82	10	0	0	0	0	2	5	13	0	47	23
824	Dipterocarpaceae	*Shorea parvistipulata* Heim	10	19.2	44.35	11	14	1	7	7	10	1	11	4	21	13
1265	Moraceae	*Ficus deltoidea* Jack subsp. *deltoidea*	10	19.2	47.17	13	19	9	0	10	9	2	6	7	20	4
947	Polygalaceae	*Xanthophyllum pedicellatum* Meijden	10	19.1	4.41	1	0	0	0	0	0	0	0	24	71	4
1071	Meliaceae	*Aglaia tomentosa* Teijsm. & Binn.	10	19.1	55.17	13	11	1	0	3	8	2	14	11	19	17
803	Dipterocarpaceae	*Shorea kudatensis* G.H.S.Wood ex Meijer	10	19	5.96	4	0	0	0	0	0	0	1	31	60	5
828	Dipterocarpaceae	*Shorea pubistyla* P.S.Ashton	10	19	3.05	0	1	0	2	0	1	0	0	8	87	1
1139	Caesalpiniaceae	*Intsia bijuga* Colebr. Kuntze	10	19	9.71	0	0	0	0	35	0	0	4	8	46	7
18	Annonaceae	*Maasia sumatrana* Miq. Mols. P.J.A.Keßler & S.H.Rogstad	10	18.9	49.76	13	19	3	0	2	7	4	9	9	20	14
357	Flacourtiaceae	*Homalium caryophyllaceum* Zoll. & Moritzi Benth.	10	18.9	22.29	2	0	0	11	14	0	0	1	28	29	14
893	Olacaceae	*Ochanostachys amentacea* Mast.	10	18.9	39.99	7	12	0	5	4	2	11	6	10	22	20
83	Euphorbiaceae	*Aporosa benthamiana* Hook.f.	10	18.8	35.18	22	22	2	0	0	3	7	6	3	24	11
669	Ulmaceae	*Trema tomentosa* Roxb. Hara	10	18.7	45.98	10	5	1	0	3	2	11	13	16	21	18
741	Dipterocarpaceae	*Dryobalanops* keithii Symington	10	18.7	8.36	6	0	0	0	0	0	0	3	28	50	13
829	Dipterocarpaceae	*Shorea quadrinervis* Slooten	10	18.7	13.79	0	14	0	0	23	17	2	0	9	36	0
415	Celastraceae	*Bhesa paniculata* Arn.	10	18.5	44.35	10	17	7	0	13	11	7	2	6	20	7
548	Fagaceae	*Castanopsis evansii* Elmer	10	18.5	32.82	4	0	0	20	15	2	2	5	16	23	14
691	Anacardiaceae	*Melanochyla castaneifolia* Ding Hou	10	18.5	18.16	7	12	0	0	0	0	6	4	23	32	16
942	Polygalaceae	*Xanthophyllum heterophyllum* Meijden	10	18.5	4.33	0	0	0	0	0	0	0	0	27	70	3
1293	Moraceae	*Ficus pisifera* Wall. ex Voigt	10	18.5	47.84	8	19	0	1	10	8	0	10	14	19	10
319	Burseraceae	*Santiria griffithii* Engl.	10	18.4	21.17	2	23	1	0	0	1	8	2	14	29	22
361	Flacourtiaceae	*Homalium panayanum* Fern.-Vill.	10	18.4	10.71	13	0	0	0	0	0	0	0	19	45	22
395	Rhizophoraceae	*Bruguiera cylindrica* Blume	10	18.4	9.01	0	0	0	11	29	0	0	1	12	45	2
509	Simaroubaceae	*Quassia indica* Gaertn. Noot.	10	18.4	10.05	0	0	0	0	12	0	0	1	19	45	22
613	Vitaceae	*Leea indica* Burm.f. Merr.	10	18.4	40.53	13	18	0	0	4	0	1	13	15	22	15
73	Euphorbiaceae	*Antidesma neurocarpum* Miq.	10	18.3	54.69	14	13	1	0	2	7	3	14	12	19	15
815	Dipterocarpaceae	*Shorea obscura* Meijer	10	18.3	35.19	14	19	1	0	0	13	4	4	16	22	7
103	Euphorbiaceae	*Baccaurea edulis* Merr.	10	18.2	27.55	20	11	0	0	0	4	7	1	7	27	23
608	Cyperaceae	*Scleria motleyi* C.B.Clarke	10	18.2	13.68	31	0	0	0	0	0	0	0	2	41	26
678	Anacardiaceae	*Drimycarpus luridus* Hook.f. Ding Hou	10	18.1	14.32	0	0	0	0	0	0	3	4	20	38	36
1014	Sapindaceae	*Arytera litoralis* Blume	10	18.1	32.59	11	0	0	0	0	2	1	14	25	25	21
1048	Meliaceae	*Aglaia crassinervia* Kurz ex Hiern	10	18.1	41.7	17	0	4	0	0	12	4	6	16	22	19
65	Euphorbiaceae	*Agrostistachys borneensis* Becc.	10	17.9	25.64	11	22	1	0	13	2	3	5	11	26	7
895	Olacaceae	*Scorodocarpus borneensis* Becc.	10	17.9	42.56	4	11	2	0	4	11	13	10	10	20	14
1287	Moraceae	*Ficus microcarpa* L.f.	10	17.9	30.91	4	0	0	7	23	3	5	1	18	24	14
42	Annonaceae	*Polyalthia clavigera* King	10	17.8	7.44	0	9	0	0	27	4	0	0	10	49	0
255	Myricaceae	*Myrica esculenta* Buch.-Ham. ex D.Don	10	17.8	11.34	23	0	0	9	1	0	0	2	22	42	0
356	Flacourtiaceae	*Flacourtia rukam* Zoll. & Moritzi	10	17.8	47.81	13	17	3	0	11	12	2	2	10	19	11
430	Celastraceae	*Microtropis kinabaluensis* Merr. & Freeman	10	17.8	11.25	27	0	2	0	3	2	0	3	13	44	6
609	Cyperaceae	*Scleria purpurascens* Steud.	10	17.8	39.79	20	15	4	0	2	0	4	7	11	22	15
1107	Caesalpiniaceae	*Bauhinia excelsa* Blume ex Miq. Prain var. *excelsa*	10	17.8	25.14	11	0	0	0	4	0	0	10	25	28	21

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
306	Burseraceae	*Dacryodes incurvata* Engl. H.J.Lam	10	17.7	50.71	12	12	10	0	0	16	2	6	15	18	9
619	Hypericaceae	*Cratoxylum arborescens* Vahl Blume var. *borneense* A.C.Church & P.F.Stevens	10	17.7	33.43	7	18	2	0	4	2	10	3	17	22	14
1083	Meliaceae	*Dysoxylum alliaceum* Blume Blume	10	17.7	50.12	11	3	3	0	1	9	12	7	16	19	19
558	Fagaceae	*Lithocarpus caudatifolius* Merr. Rehder	10	17.6	14.07	19	0	0	0	0	0	0	0	15	39	27
585	Passifloraceae	*Adenia macrophylla* Blume Koord. var. *macrophylla*	10	17.6	50.11	4	12	0	0	8	8	11	6	17	18	16
735	Dipterocarpaceae	*Dipterocarpus stellatus* Vesque subsp. *parvus*	10	17.6	8.5	0	1	0	0	0	0	8	0	4	49	38
1211	Myristicaceae	*Knema latifolia* Warb.	10	17.6	46.4	12	14	4	0	8	2	11	3	7	19	19
1264	Moraceae	*Ficus delosyce* Corner var. *obtusa* Corner	10	17.6	16.84	0	0	0	0	1	1	0	21	27	35	16
724	Dipterocarpaceae	*Dipterocarpus gracilis* Blume	10	17.5	32.14	17	2	0	0	10	2	2	5	21	24	17
22	Annonaceae	*Mezzettia parviflora* Becc.	10	17.4	31.67	2	7	0	8	11	1	15	0	19	22	16
393	Proteaceae	*Helicia robusta* Roxb. R.Br. ex Wall.	10	17.4	9.64	32	3	0	0	0	0	3	3	0	48	11
676	Anacardiaceae	*Campnosperma coriaceum* Jack Hallier f.	10	17.4	6.53	0	0	0	0	27	0	0	6	6	55	5
850	Dipterocarpaceae	*Vatica globosa* P.S.Ashton	10	17.4	13.02	0	32	0	0	10	4	1	13	4	36	2
1022	Sapindaceae	*Guioa pleuropteris* Blume Radlk.	10	17.4	13.51	20	0	0	0	0	0	0	6	8	40	25
1100	Meliaceae	*Vavaea amicorum* Benth.	10	17.4	13.08	7	0	0	0	0	0	0	31	1	42	19
1419	Orchidaceae	*Liparis lacerata* Ridl.	10	17.3	15.89	7	19	0	0	1	0	8	8	1	34	23
204	Combretaceae	*Terminalia subspathulata* King	10	17.2	23.43	0	0	0	13	21	1	0	14	23	27	1
705	Anacardiaceae	*Swintonia schwenckii* Teijsm. & Binn. Teijsm. & Binn. ex Hook.f.	10	17.2	27.63	0	15	0	18	2	9	6	2	11	23	13
202	Combretaceae	*Terminalia foetidissima* Griff.	10	17.1	11.12	0	0	0	0	17	0	4	0	5	41	33
335	Connaraceae	*Connarus semidecandrus* Jack	10	17.1	27.54	0	9	0	6	23	0	0	2	24	24	12
519	Icacinaceae	*Gonocaryum macrophyllum* Blume Sleumer	10	17.1	9.92	9	0	0	0	4	0	0	0	1	46	40
912	Ixonanthaceae	*Ixonanthes petiolaris* Blume	10	17.1	36.49	4	13	0	0	0	17	15	5	9	21	14
971	Chrysobalaceae	*Parinari elmeri* Merr.	10	17.1	21.8	1	18	0	0	0	4	8	7	9	28	24
171	Euphorbiaceae	*Microstachys chamaelea* L. Müll.Arg.	10	17	12.89	0	0	0	0	18	0	1	2	27	37	14
270	Thymelaeaceae	*Gonystylus maingayi* Hook.f.	10	17	7.47	0	14	0	0	20	1	3	4	9	48	1
590	Cyperaceae	*Cyperus cyperinus* Retz. J.V.Suringar	10	17	27.64	15	19	1	0	9	0	0	20	1	26	9
734	Dipterocarpaceae	*Dipterocarpus sarawakensis* Slooten	10	17	11.13	0	35	0	23	0	3	0	1	1	35	2
36	Annonaceae	*Phaeanthus ebracteolatus* C.Presl Merr.	10	16.9	35.04	21	8	7	0	0	6	4	8	8	23	15
253	Gnetaceae	*Gnetum leptostachyum* Blume var. *robustum* Markgr.	10	16.9	38.91	11	19	12	2	11	17	1	3	2	20	3
1172	Loranthaceae	*Lepeostegeres lanceifolius* Tiegh. Danser	10	16.8	14.96	6	0	0	0	0	6	0	2	19	36	30
208	Convolvulaceae	*Erycibe impressa* Hoogland	10	16.7	29.35	24	0	1	0	0	0	1	11	21	26	17
518	Icacinaceae	*Gonocaryum cognatum* Elmer	10	16.6	14.71	4	7	0	0	0	1	1	3	26	35	23
624	Bignoniaceae	*Deplanchea bancana* Scheff. Steenis	10	16.6	15.19	33	24	1	0	2	0	4	0	1	34	1
1163	Loranthaceae	*Dendrophthoe pentandra* L. Miq.	10	16.5	38	10	2	9	12	14	6	0	5	17	20	6
436	Epacridaceae	*Styphelia malayana* Jack Spreng.	10	16.4	3.72	9	7	0	0	0	0	0	0	12	72	0
763	Dipterocarpaceae	*Hopea tenuinervula* P.S.Ashton	10	16.4	16.09	0	4	0	0	0	3	2	26	27	34	4
1064	Meliaceae	*Aglaia palembanica* Miq.	10	16.4	6.06	14	0	0	0	0	1	0	14	1	59	11
120	Euphorbiaceae	*Elateriospermum tapos* Blume	10	16.3	45.4	7	18	7	3	11	9	6	4	6	18	13
354	Flacourtiaceae	*Casearia tuberculata* Blume	10	16.3	21.37	24	0	5	0	0	16	3	13	0	30	9
633	Crypteroniaceae	*Dactylocladus stenostachys* Oliv.	10	16.3	16.67	0	0	0	0	29	5	2	1	29	31	2
739	Dipterocarpaceae	*Dryobalanops beccarii* Dyer	10	16.2	46.19	2	9	1	8	10	6	13	2	14	18	18
1076	Meliaceae	*Chisocheton patens* Blume	10	16.2	38.41	10	4	0	3	10	4	2	15	14	21	16
137	Euphorbiaceae	*Macaranga didymocarpa* Whitmore	10	16.1	12.44	0	15	0	0	14	9	5	4	18	35	0
823	Dipterocarpaceae	*Shorea parvifolia* Dyer subsp. *velutinata* P.S.Ashton	10	16.1	58.47	9	15	0	1	13	5	8	8	10	16	14
1111	Caesalpiniaceae	*Bauhinia havilandii* Merr.	10	16.1	6.7	2	9	0	0	0	0	1	2	2	53	32
140	Euphorbiaceae	*Macaranga indistincta* Whitmore	10	16	9.01	32	0	0	0	0	0	0	3	0	48	17
1234	Nepenthaceae	*Nepenthes ampullaria* Jack	10	16	57.35	5	8	1	14	12	11	3	6	15	16	8
133	Euphorbiaceae	*Macaranga caladiifolia* Becc.	10	15.9	19.49	0	15	0	2	27	10	6	2	10	27	0
136	Euphorbiaceae	*Macaranga depressa* Müll.Arg. Müll.Arg.	10	15.9	6.69	22	0	0	0	0	0	0	5	16	54	2
1160	Loranthaceae	*Dendrophthoe constricta* Korth. Danser	10	15.9	25.65	12	0	0	0	6	3	2	3	23	26	26

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
41	Annonaceae	*Polyalthia cauliflora* Hook.f. & Thomson var. *desmantha* Hook.f. & Thomson J.Sinclair	10	15.8	18.61	8	23	1	0	1	13	0	15	1	30	9
617	Hypericaceae	*Cratoxylum arborescens* Vahl Blume	10	15.8	36.24	13	5	8	0	0	6	15	3	12	21	18
506	Simaroubaceae	*Brucea javanica* L. Merr.	10	15.7	29	10	13	0	0	0	1	2	10	17	24	23
946	Polygalaceae	*Xanthophyllum nigricans* Meijden	10	15.7	16.58	3	16	0	2	0	6	8	11	1	32	22
147	Euphorbiaceae	*Macaranga praestans* Airy Shaw	10	15.6	14.22	13	23	0	0	20	0	1	1	6	33	4
1091	Meliaceae	*Dysoxylum rugulosum* King	10	15.6	41.19	12	2	7	0	0	15	13	5	11	20	15
953	Polygalaceae	*Xanthophyllum stipitatum* A.W.Benn.	10	15.5	61.11	12	14	1	0	5	10	10	5	12	16	15
1171	Loranthaceae	*Lepeostegeres centiflorus* Stapf Tiegh.	10	15.5	4.84	34	0	0	0	0	0	0	0	1	65	0
213	Convolvulaceae	*Merremia clemensiana* Ooststr.	10	15.4	13.16	10	13	0	0	0	3	1	2	4	36	30
527	Icacinaceae	*Stemonurus scorpioides* Becc.	10	15.4	10.38	0	0	0	0	27	0	0	0	29	39	5
995	Mimosaceae	*Parkia singularis* Miq.	10	15.4	26.19	9	5	11	1	2	3	19	10	8	24	8
1002	Rosaceae	*Prunus javanica* Teijsm. & Binn. Miq.	10	15.4	36.11	7	1	1	17	6	8	7	5	11	21	17
1084	Meliaceae	*Dysoxylum arborescens* Blume Miq.	10	15.2	26.49	13	0	5	0	7	1	2	10	17	25	19
682	Anacardiaceae	*Gluta sabahana* Ding Hou	10	15.1	12.93	20	0	0	0	0	0	0	7	15	38	20
701	Anacardiaceae	*Semecarpus glaucus* Engl.	10	15.1	32.84	1	19	1	2	13	18	5	5	10	20	6
856	Dipterocarpaceae	*Vatica nitens* King	10	15	12.37	0	30	0	1	7	5	8	0	6	33	11
1080	Meliaceae	*Chisocheton polyandrus* Merr.	10	15	16.43	21	0	0	0	0	0	0	19	14	34	12
981	Mimosaceae	*Albizia rosulata* Kosterm. I.C.Nielsen	10	14.9	15.46	1	0	0	0	21	1	1	5	24	32	16
1226	Myristicaceae	*Myristica borneensis* Warb.	10	14.9	22.65	3	24	3	2	9	23	6	4	3	24	0
1029	Sapindaceae	*Lepisanthes multijuga* Hook.f. Leenh.	10	14.8	4.1	22	0	0	0	0	0	0	0	0	68	10
262	Thymelaeaceae	*Amyxa pluricornis* Radlk. Domke	10	14.6	28.37	0	19	5	0	9	19	12	5	9	21	1
1077	Meliaceae	*Chisocheton pentandrus* Blanco Merr.	10	14.5	7.75	22	0	0	0	0	0	0	4	0	49	25
185	Euphorbiaceae	*Syndyophyllum occidentale* Airy Shaw Welzen	10	14.4	14.67	13	0	0	0	0	0	0	11	21	34	21
533	Oxalidaceae	*Dapania grandifolia* Veldkamp	10	14.4	9.47	1	0	0	0	0	0	0	1	14	42	42
635	Lentibulariaceae	*Utricularia minutissima* Vahl	10	14.3	10.29	0	0	0	22	2	1	3	4	23	37	8
1300	Moraceae	*Ficus septica* Burm.f.	10	14.1	13.93	31	0	0	0	0	0	0	3	7	36	24
201	Combretaceae	*Terminalia citrina* Gaertn. Roxb. ex Fleming	10	14	9.01	3	0	0	0	0	0	0	0	12	43	43
821	Dipterocarpaceae	*Shorea pachyphylla* Ridl. ex Symington	10	14	15.93	0	4	0	1	15	0	17	3	18	29	12
689	Anacardiaceae	*Mangifera rufocostata* Kosterm.	10	13.9	13.86	16	0	0	0	0	0	0	0	19	34	30
718	Dipterocarpaceae	*Dipterocarpus conformis* Slooten subsp. *borneensis* P.S.Ashton	10	13.8	15.19	33	9	0	0	0	0	0	0	0	33	24
30	Annonaceae	*Neo-uvaria foetida* Maingay ex Hook.f. & Thomson Airy Shaw	10	13.7	9.84	12	0	0	0	0	0	0	3	5	41	40
78	Euphorbiaceae	*Antidesma puncticulatum* Miq.	10	13.7	8.03	0	0	0	0	8	0	0	0	31	43	18
349	Flacourtiaceae	*Casearia capitellata* Blume	10	13.7	15.53	26	0	0	0	0	0	0	0	19	32	22
766	Dipterocarpaceae	*Hopea wyattsmithii* G.H.S.Wood ex P.S.Ashton	10	13.7	1.85	0	0	0	0	0	0	0	0	2	97	1
699	Anacardiaceae	*Semecarpus bunburyanus* Gibbs	10	13.6	15.32	27	0	0	0	0	0	0	5	8	33	26
983	Mimosaceae	*Archidendron borneense* Benth. I.C.Nielsen	10	13.6	22.71	5	2	0	5	21	0	1	9	21	25	10
1429	Orchidaceae	*Podochilus lucescens* Blume	10	13.5	27.07	24	1	6	0	9	8	2	12	3	24	12
31	Annonaceae	*Orophea alba* P.J.A.Kessler	10	13.4	6.48	3	0	0	0	0	0	0	2	45	47	3
383	Goodeniaceae	*Scaevola sericea* Vahl	10	13.4	2.9	0	0	0	0	0	0	0	0	9	75	16
405	Rhizophoraceae	*Kandelia candel* L. Druce	10	13.4	7.04	0	0	0	0	23	0	0	2	24	45	6
427	Celastraceae	*Lophopetalum sessilifolium* Ridl.	10	13.4	15.67	0	27	0	0	0	1	0	29	13	30	0
167	Euphorbiaceae	*Mallotus philippensis* (Lam.) Műll.Arg.	10	13.2	5.64	20	0	0	0	0	0	0	0	11	54	15
259	Sonneratiaceae	*Sonneratia ovata* Backer	10	13.2	9.54	0	0	0	0	16	0	1	2	24	38	18
1096	Meliaceae	*Reinwardtiodendron kinabaluense* Kosterm. Mabb.	10	13.1	12.18	36	0	0	0	0	0	0	4	14	37	10
186	Euphorbiaceae	*Triadica cochinchinensis* Lour.	10	13	6.56	25	0	0	0	4	0	0	6	15	49	0
725	Dipterocarpaceae	*Dipterocarpus grandiflorus* Blanco	10	12.9	5.85	0	0	0	0	0	0	0	0	8	51	41
990	Mimosaceae	*Archidendron kunstleri* Prain I.C.Nielsen subsp. *ashtonii* Nielsen	10	12.9	1.74	0	0	0	0	0	0	0	0	0	97	3
505	Simaroubaceae	*Allantospermum borneense* Forman	10	12.8	12.74	1	5	0	23	0	0	1	1	9	32	28
814	Dipterocarpaceae	*Shorea obovoidea* Slooten	10	12.8	6.06	0	36	0	0	0	0	1	16	0	46	0
104	Euphorbiaceae	*Baccaurea motleyana* Müll.Arg. Müll.Arg.	10	12.7	17.98	2	21	0	0	7	2	1	4	24	26	14
638	Symplocaceae	*Symplocos adenophylla* Wall. & G.Don var. *adenophylla* Wall. ex G.Don	10	12.7	22.11	20	4	5	0	15	10	0	7	8	25	7
923	Magnoliaceae	*Michelia montana* Blume	10	12.7	12.15	8	0	0	0	0	0	0	5	26	34	27

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
19	Annonaceae	*Marsypopetalum pallidum* Blume Backer	10	12.6	7.56	23	0	0	0	0	0	0	0	0	46	31
1261	Moraceae	*Ficus brunneo-aurata* Corner	10	12.6	19.97	9	21	2	0	10	16	0	13	2	25	1
786	Dipterocarpaceae	*Shorea exelliptica* Meijer	10	12.5	25.74	1	16	0	2	8	4	10	8	10	22	21
1252	Moraceae	*Artocarpus tomentosulus* Jarrett	10	12.5	11.66	24	0	0	0	0	2	5	15	0	37	17
1233	Nepenthaceae	*Nepenthes albo-marginata* Lobb ex Lindl.	10	12.4	11.11	0	4	0	0	33	21	1	2	5	33	0
419	Celastraceae	*Kokoona littoralis* M.A.Lawson	10	12.3	17.21	0	17	0	6	14	7	12	1	16	25	2
532	Ochnaceae	*Schuurmansiella angustifolia* Hook.f. Hallier f.	10	12.1	13.84	0	26	0	1	18	6	2	3	17	27	0
1384	Orchidaceae	*Coelogyne verrucosa* S.E.C.Sierra	10	12.1	12.39	29	6	0	0	0	0	2	1	0	35	28
293	Burseraceae	*Canarium hirsutum* Willd. f. scabrum Blume Leenh.	10	12	25.37	11	0	1	0	21	10	0	9	17	23	8
485	Ericaceae	*Vaccinium lobbii* Ridl. Sleumer	10	11.8	2.11	0	0	0	0	0	18	0	0	0	82	0
67	Euphorbiaceae	*Antidesma brachybotrys* Airy Shaw	10	11.4	13.75	17	16	0	0	8	0	3	4	1	30	22
959	Sabiaceae	*Meliosma lanceolata* Blume	10	11.4	6.14	0	0	0	5	20	0	0	1	30	43	0
614	Taccaceae	*Tacca bibracteata* Drenth	10	11.3	14.83	20	17	0	0	0	0	4	1	1	29	27
752	Dipterocarpaceae	*Hopea griffithii* Kurz	10	11.3	11.66	0	12	0	0	24	14	4	5	11	30	0
1276	Moraceae	*Ficus gul* Lauterb. & K.Schum.	10	11.2	11.72	19	0	0	0	0	0	1	8	8	34	30
2	Annonaceae	*Alphonsea kinabaluensis* J.Sinclair	10	11.1	9.05	25	31	1	0	0	2	0	2	4	35	0
231	Dilleniaceae	*Saurauia longistyla* Merr.	10	11.1	4.31	29	0	0	0	0	0	0	0	0	58	13
683	Anacardiaceae	*Gluta velutina* Blume	10	10.9	11.25	0	0	0	0	22	0	0	1	30	31	15
1386	Orchidaceae	*Dendrobium aloifolium* Blume Rchb.f.	10	10.9	17.78	10	13	4	0	0	1	6	0	22	25	20
141	Euphorbiaceae	*Macaranga kinabaluensis* Airy Shaw	10	10.8	11.3	34	0	0	0	0	4	0	0	1	35	27
762	Dipterocarpaceae	*Hopea semicuneata* Symington	10	10.5	4.92	9	0	0	0	0	0	0	9	19	50	13
1008	Rosaceae	*Rubus fraxinifolius* Poir.	10	10.5	9.76	36	0	4	0	0	2	0	4	14	36	5
69	Euphorbiaceae	*Antidesma ghaesembilla* Gaertn.	10	10.4	15.81	9	0	0	0	10	2	0	12	18	27	22
1112	Caesalpiniaceae	*Bauhinia integrifolia* Roxb. subsp. *cumingiana* Benth.	10	10.4	9.91	18	0	0	0	0	0	0	1	18	35	28
43	Annonaceae	*Polyalthia congesta* Ridl. J.Sinclair	10	10.3	9.75	2	0	0	0	13	0	0	7	13	35	29
1121	Caesalpiniaceae	*Bauhinia wrayi* Prain ex King var. *cardiohylla* Merr. K.Larsen & S.S.Larsen	10	10.3	3.33	0	0	0	0	18	0	0	2	21	58	1
493	Loganiaceae	*Fagraea blumei* G.Don subsp. *blumei* G.Don	10	10.2	18.75	18	22	10	1	9	7	1	0	8	22	1
1208	Myristicaceae	*Knema kunstleri* King Warb. subsp. *coriacea* Warb. W.J.de Wilde	10	9.9	4.33	0	0	0	0	39	1	0	0	11	49	0
152	Euphorbiaceae	*Macaranga trichocarpa* Rchb.f. & Zoll. Müll.Arg.	10	9.8	27.66	8	0	0	0	17	5	3	13	15	20	18
931	Podocarpaceae	*Podocarpus micropedunculatus* de Laub.	10	9.8	1.48	0	0	0	0	0	0	0	0	0	91	9
686	Anacardiaceae	*Mangifera caesia* Jack	10	9.6	18.76	5	0	0	0	21	0	0	6	22	23	22
877	Araliaceae	*Schefflera longifructescens* Elmer	10	9.6	17.41	17	0	0	0	9	1	0	21	0	26	25
522	Icacinaceae	*Iodes philippinensis* Merr.	10	9.1	14.67	26	0	1	0	0	0	0	4	22	27	20
313	Burseraceae	*Dacryodes rostrata* Blume H.J.Lam f. pubescens Lam.	10	8.7	2.51	6	0	0	0	0	0	0	0	7	65	21
1166	Loranthaceae	*Helixanthera maxwelliana* Gibbs Danser	10	8.3	7.87	27	0	0	0	0	0	0	0	11	36	26
1134	Caesalpiniaceae	*Cynometra mirabilis* Meeuwen	10	8	4.8	2	0	0	0	0	0	0	16	0	46	35
529	Ochnaceae	*Brackenridgea palustris* Bartl.	10	7.8	4.72	0	0	0	0	17	11	0	0	22	41	9
337	Connaraceae	*Ellipanthus tomentosus* Kurz	10	7.6	2.29	23	0	0	0	0	0	0	0	0	66	11
842	Dipterocarpaceae	*Shorea symingtonii* G.H.S.Wood	10	7.6	6.73	0	0	0	0	0	0	0	34	1	38	27
1019	Sapindaceae	*Dimocarpus longan* Lour. var. *echinatus* Leenh.	10	7.5	11.03	19	0	0	0	0	0	0	5	22	28	26
1198	Myristicaceae	*Horsfieldia oligocarpa* Warb.	10	7.5	1.01	0	0	0	0	0	0	0	0	4	96	0
1127	Caesalpiniaceae	*Caesalpinia parviflora* Prain ex King	10	7.4	12.23	15	0	0	0	0	0	0	7	26	26	26
1203	Myristicaceae	*Horsfieldia sterilis* W.J.de Wilde	10	7.4	1.52	6	0	0	0	0	0	0	1	0	79	14
381	Flacourtiaceae	*Xylosma luzonensis* Presl Clos	10	7.3	5.36	29	0	0	0	0	0	0	0	6	41	23
158	Euphorbiaceae	*Mallotus havilandii* Airy Shaw	10	7.1	2.61	0	1	0	0	0	18	0	2	25	54	0
1110	Caesalpiniaceae	*Bauhinia foraminifera* Gagnep.	10	7	1.54	0	0	0	0	3	0	0	1	24	72	0
1320	Orchidaceae	*Appendicula buxifolia* Blume	10	7	3.19	16	0	0	0	0	0	0	0	0	52	32
437	Epacridaceae	*Styphelia suaveolens* Hook.f. Warb.	10	6.8	8.11	32	0	0	0	0	0	0	13	5	33	17
168	Euphorbiaceae	*Mallotus tiliifolius* Blume Müll.Arg.	10	5.9	1.29	0	0	0	0	15	0	0	0	13	72	0
831	Dipterocarpaceae	*Shorea revoluta* P.S.Ashton	10	5.8	0.73	0	0	0	0	0	0	0	0	0	100	0
1049	Meliaceae	*Aglaia cumingiana* Turcz.	10	5.7	3.86	40	0	0	0	0	0	0	0	8	43	8

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
521	Icacinaceae	*Iodes cirrhosa* Turcz.	10	5.6	9.33	20	0	0	0	0	0	0	25	7	28	21
631	Cornaceae	*Mastixia trichotoma* Blume var. *maingayi* Clarke Danser	10	4.8	4.96	18	0	0	0	0	0	18	0	1	33	29
854	Dipterocarpaceae	*Vatica maritima* Slooten	10	4	0.56	0	0	0	0	0	0	0	4	0	96	0
1081	Meliaceae	*Chisocheton ruber* Ridl.	10	3.6	0.51	0	0	0	0	0	0	0	0	7	93	0
564	Fagaceae	*Lithocarpus echinulatus* Soepadmo	11	41.7	15.09	4	0	0	0	5	0	3	4	17	11	55
976	Mimosaceae	*Adenanthera kostermansii* I.C.Nielsen	11	41.7	10.26	1	0	0	0	6	0	6	4	3	12	68
60	Annonaceae	*Uvaria elmeri* Merr.	11	39.2	8.58	13	0	0	0	0	0	1	5	3	3	74
1246	Moraceae	*Artocarpus lanceifolius* Roxb.	11	38.5	17.54	18	10	0	0	0	0	0	3	2	16	50
504	Simaroubaceae	*Ailanthus integrifolia* Lam.	11	37.6	21.87	1	0	0	0	3	0	19	4	8	22	43
1136	Caesalpiniaceae	*Dialium hydnocarpoides* de Wit	11	37.6	13.19	1	0	0	0	4	0	2	15	6	15	58
54	Annonaceae	*Popowia hirta* Miq.	11	35.8	17.59	8	0	0	0	5	0	5	1	7	25	48
211	Convolvulaceae	*Ipomoea triloba* L.	11	35.4	14.57	1	0	0	0	4	0	18	1	6	19	51
685	Anacardiaceae	*Koordersiodendron pinnatum* Blanco Merr.	11	35.2	27.26	14	0	0	0	0	0	5	3	16	23	39
1151	Caesalpiniaceae	*Uittienia modesta* Steenis	11	34.5	5.98	1	0	0	0	0	0	0	1	0	15	83
106	Euphorbiaceae	*Baccaurea parviflora* Mülll.Arg. Mülll.Arg.	11	32.6	30.72	13	0	0	7	0	0	5	5	7	27	35
1131	Caesalpiniaceae	*Crudia subsimplicifolia* Merr.	11	32	12.69	16	0	0	0	0	0	0	0	9	21	55
420	Celastraceae	*Kokoona ochracea* Elmer Merr.	11	31.8	17.93	17	1	0	0	0	0	0	0	2	33	47
1055	Meliaceae	*Aglaia laxiflora* Miq.	11	30.9	23.14	17	0	0	0	9	0	3	4	1	26	40
726	Dipterocarpaceae	*Dipterocarpus humeratus* Sloot.	11	30.8	26.91	3	6	1	0	0	0	17	3	8	26	35
1132	Caesalpiniaceae	*Crudia tenuipes* Merr.	11	30.8	29.64	16	7	0	0	0	0	2	1	7	32	35
38	Annonaceae	*Polyalthia borneensis* Merr.	11	30.5	13.35	23	0	0	0	0	1	1	1	1	20	53
702	Anacardiaceae	*Spondias philippinensis* Elmer Airy Shaw & Forman	11	30.5	27.14	8	0	0	0	0	0	3	1	26	27	35
1130	Caesalpiniaceae	*Crudia reticulata* Merr.	11	30.2	19.21	12	2	0	0	0	0	0	8	12	23	43
1015	Sapindaceae	*Dictyoneura acuminata* Blume subsp. *acuminata*	11	29.3	13.87	13	0	0	1	4	0	22	5	4	4	48
308	Burseraceae	*Dacryodes macrocarpa* King H.J.Lam var. *kostermansii* Kalkman	11	28.2	20.36	1	0	0	15	4	2	13	3	8	17	37
62	Annonaceae	*Uvaria ovalifolia* Blume var. *borneensis* Miq.	11	27.9	28.11	15	0	1	0	1	0	21	1	13	17	33
674	Anacardiaceae	*Buchanania insignis* Blume	11	27.9	13.84	10	0	0	0	0	0	10	3	14	16	47
143	Euphorbiaceae	*Macaranga lowii* King ex Hook.f. var. *lowii*	11	27.7	25.59	7	0	0	25	0	0	10	9	1	14	33
48	Annonaceae	*Polyalthia microtus* Miq.	11	27.3	32.09	7	0	0	3	9	0	4	5	13	28	31
1034	Sapindaceae	*Nephelium cuspidatum* Blume var. *cuspidatum*	11	26.9	33.53	5	0	0	0	0	5	5	21	21	13	30
1141	Caesalpiniaceae	*Koompassia excelsa* Becc. Taub.	11	26.6	17.7	7	0	0	0	0	0	1	4	7	38	43
330	Connaraceae	*Agelaea trinervis* Llanos Merr.	11	26.5	18.01	4	0	0	0	0	0	0	3	13	38	42
863	Dipterocarpaceae	*Vatica umbonata* Hook.f. Burck subsp. *acrocarpa* Slooten P.S.Ashton	11	26.2	13.75	37	0	0	0	0	0	0	0	0	14	49
59	Annonaceae	*Sageraea lanceolata* Miq.	11	26.1	21.38	4	0	0	0	8	0	0	3	15	32	37
117	Euphorbiaceae	*Chaetocarpus castanocarpus* Roxb. Thwaites	11	26.1	41.35	9	4	0	0	7	1	7	9	10	26	27
517	Icacinaceae	*Gonocaryum calleryanum* Baill. Becc.	11	25.8	14.15	0	0	0	12	4	0	15	2	19	7	42
934	Polygalaceae	*Polygala paniculata* L.	11	25.8	13.52	20	0	0	0	14	0	0	0	6	12	47
24	Annonaceae	*Miliusa macropoda* Miq.	11	25.2	11.02	4	0	0	0	0	0	0	15	17	13	52
299	Burseraceae	*Canarium megalanthum* Merr.	11	25	13.43	0	0	0	0	4	0	8	2	12	29	45
291	Burseraceae	*Canarium denticulatum* Blume subsp. *kostermansii* Leenh.	11	24.9	10.62	19	0	0	0	0	0	0	3	4	20	54
424	Celastraceae	*Lophopetalum javanicum* Zoll. Turcz.	11	24.8	39.93	12	0	0	3	7	0	4	12	16	18	26
350	Flacourtiaceae	*Casearia grewiaefolia* Vent.	11	24.4	18	2	0	0	15	1	0	3	19	15	8	38
778	Dipterocarpaceae	*Shorea bracteolata* Dyer	11	24.3	32.61	6	22	2	0	1	0	2	7	6	27	28
970	Chrysobalaceae	*Parinari canarioides* Kosterm.	11	23.9	18.85	0	1	0	0	0	0	25	0	18	21	36
769	Dipterocarpaceae	*Shorea agami* P.S.Ashton subsp. *agami* P.S.Ashton	11	23.7	16.84	1	1	0	0	0	0	5	1	22	31	39
796	Dipterocarpaceae	*Shorea guiso* Blanco Blume	11	23.7	24.37	8	0	0	0	0	0	0	1	32	25	33
407	Rhizophoraceae	*Pellacalyx lobbii* Hook.f. Schimp.	11	23.5	25.6	5	2	0	0	0	10	4	1	23	24	31
1149	Caesalpiniaceae	*Sindora wallichii* Benth.	11	23.5	9.3	2	0	0	0	0	0	1	1	5	36	55
312	Burseraceae	*Dacryodes rostrata* Blume H.J.Lam f. pallida H.J.Lam	11	23.3	16.56	0	0	0	0	3	0	7	3	24	24	39
794	Dipterocarpaceae	*Shorea gibbosa* Bandis	11	23.3	42.78	13	19	1	0	0	4	4	6	8	21	24
876	Araliaceae	*Schefflera filipes* Merr.	11	23.3	18.19	31	0	0	0	0	0	0	13	2	14	41

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
111	Euphorbiaceae	*Baccaurea tetrandra* Baill. Mülll.Arg.	11	23.1	46.17	12	14	1	0	1	0	5	14	10	20	23
1009	Rosaceae	*Rubus moluccanus* L.	11	23.1	14.35	4	0	0	0	2	0	6	11	30	5	41
145	Euphorbiaceae	*Macaranga pearsonii* Merr.	11	23	26.91	14	0	0	0	1	0	4	7	21	21	31
1222	Myristicaceae	*Knema pubiflora* W.J.de Wilde	11	23	26.85	0	4	0	0	0	1	12	3	24	26	30
507	Simaroubaceae	*Eurycoma longifolia* Jack	11	22.9	37.86	18	9	2	0	0	0	11	1	11	23	25
958	Polygalaceae	*Xanthophyllum vitellinum* Blume D.Dietr.	11	22.7	39.37	9	0	0	0	0	9	9	11	13	23	25
1188	Myristicaceae	*Horsfieldia borneensis* W.J.de Wilde	11	22.5	12.85	0	12	0	0	2	0	2	0	7	34	43
972	Chrysobalaceae	*Parinari oblongifolia* Hook.f.	11	22.4	41.03	12	0	0	0	0	10	4	13	20	16	25
25	Annonaceae	*Mitrephora fragrans* Merr.	11	22.2	12.18	1	0	0	0	0	0	0	7	33	15	44
53	Annonaceae	*Popowia bancana* Scheff.	11	22.2	14.59	10	0	1	0	5	1	24	7	5	7	40
1129	Caesalpiniaceae	*Crudia ornata* de Wit	11	21.8	34.38	12	0	0	0	0	1	2	17	22	18	27
656	Symplocaceae	*Symplocos odoratissima* Choisy ex Zoll. var. *wenzelii* Merr. Noot.	11	21.1	9.83	7	0	0	0	0	0	0	0	6	37	51
1286	Moraceae	*Ficus melinocarpa* Blume	11	20.6	12.01	15	0	0	0	0	0	0	6	6	27	46
719	Dipterocarpaceae	*Dipterocarpus cornutus* Dyer	11	20.4	12.7	0	0	0	0	1	0	3	12	35	8	41
1137	Caesalpiniaceae	*Dialium indum* L. var. *indum*	11	20.4	49.11	16	11	3	0	3	5	9	1	15	18	20
672	Anacardiaceae	*Bouea oppositifolia* Roxb. Meisn.	11	20.3	24.78	1	0	0	11	0	0	4	10	24	21	29
1095	Meliaceae	*Reinwardtiodendron humile* Hassk. Mabb.	11	20.2	23.85	7	0	0	0	0	0	4	18	9	31	32
557	Fagaceae	*Lithocarpus bullatus* Hatus. ex Soepadmo	11	20.1	21.41	28	0	7	0	0	2	5	4	6	16	33
33	Annonaceae	*Orophea creaghii* Ridl. Leonardia & P.J.A.Kessler	11	20	17.02	14	0	0	0	0	1	0	2	12	33	37
333	Connaraceae	*Connarus grandis* Jack	11	20	19.12	16	0	0	0	0	0	0	6	11	30	36
1148	Caesalpiniaceae	*Sindora leiocarpa* Backer ex de Wit	11	19.7	36.21	3	3	0	0	20	2	11	8	10	18	24
1016	Sapindaceae	*Dimocarpus dentatus* Meijer ex Leenh.	11	19.6	11.58	16	0	0	0	0	0	0	0	10	29	45
1124	Caesalpiniaceae	*Caesalpinia latisiliqua* Cav. Hattink	11	19.4	33.05	13	20	1	0	0	0	8	2	10	21	24
986	Mimosaceae	*Archidendron clypearia* Jack I.C.Nielsen var. *clypearia*	11	19.3	12.27	0	0	0	0	5	0	8	2	18	26	41
550	Fagaceae	*Castanopsis hypophoenicea* Seemen Soepadmo	11	18.7	16.68	23	0	0	0	0	0	0	3	1	36	38
1138	Caesalpiniaceae	*Dialium maingayi* Baker	11	18.3	38.21	1	2	0	6	12	0	9	7	21	21	22
416	Celastraceae	*Bhesa robusta* Roxb. Ding Hou	11	18.2	17.13	0	1	0	0	7	0	19	12	2	25	34
275	Thymelaeaceae	*Phaleria capitata* Jack	11	18.1	17.13	19	0	0	0	15	0	1	5	9	16	35
1206	Myristicaceae	*Knema casearioides* Kosterm. W.J.de Wilde	11	18.1	10.99	28	0	0	0	0	0	0	1	0	26	46
982	Mimosaceae	*Albizia splendens* Miq.	11	18	7.37	1	0	0	0	8	0	1	0	2	34	53
996	Amaryllidaceae	*Crinum asiaticum* L.	11	18	16.47	0	2	0	0	6	0	22	4	7	24	34
174	Euphorbiaceae	*Neoscortechinia nicobarica* Hook.f. Pax & K.Hoffm.	11	17.9	12.5	2	0	0	0	0	0	0	6	16	34	41
664	Ulmaceae	*Gironniera nervosa* Planch.	11	17.8	50.68	9	12	1	2	5	5	9	11	10	18	19
818	Dipterocarpaceae	*Shorea ovalis* Korth. Blume subsp. *ovalis*	11	17.7	31.22	1	1	0	0	22	6	5	2	15	24	24
760	Dipterocarpaceae	*Hopea rudiformis* P.S.Ashton	11	17.3	2.91	0	0	0	0	0	0	0	1	1	15	84
677	Anacardiaceae	*Dracontomelon dao* Blanco Merr. & Rolfe	11	17.2	54.04	9	14	2	5	8	3	10	7	10	17	18
1133	Caesalpiniaceae	*Cynometra elmeri* Merr.	11	17.1	5.35	13	0	0	0	0	0	0	0	4	21	62
384	Malpighiaceae	*Aspidopterys elliptica* Blume Juss.	11	16.9	19.87	0	1	0	0	13	0	24	8	4	21	30
980	Mimosaceae	*Albizia pedicellata* Baker ex Benth.	11	16	15.61	16	0	0	0	3	0	8	4	4	30	35
144	Euphorbiaceae	*Macaranga motleyana* Mülll.Arg. Mülll.Arg.	11	15.9	20.68	10	0	0	0	0	2	4	15	16	23	30
336	Connaraceae	*Ellipanthus beccarii* Pierre var. *peltatus* Schellenb. Leenh.	11	15.8	14.71	23	0	0	0	0	0	0	8	2	31	37
969	Chrysobalaceae	*Maranthes corymbosa* Blume	11	15.8	16.64	11	0	0	0	5	0	0	5	13	32	33
237	Dilleniaceae	*Tetracera macrophylla* Wall. ex Hook.f. & Thomson	11	15.4	10.82	0	25	0	0	0	0	3	1	3	30	38
822	Dipterocarpaceae	*Shorea parvifolia* Dyer subsp. *parvifolia*	11	15.4	44.57	5	15	1	0	0	12	17	9	11	12	18
938	Polygalaceae	*Xanthophyllum affine* Korth.	11	15.3	6.48	0	0	0	0	5	0	0	4	34	8	50
1417	Orchidaceae	*Galeola nudifolia* Lour.	11	15.1	14.5	23	0	0	0	0	0	0	15	3	23	36
164	Euphorbiaceae	*Mallotus mollisimus* (Geiseler) Airy Shaw	11	15	28.05	15	0	0	0	0	0	6	11	20	24	25
804	Dipterocarpaceae	*Shorea laevis* Ridl.	11	14.7	53.64	5	13	1	12	2	3	11	10	12	15	16
622	Hypericaceae	*Cratoxylum sumatranum* Jack Blume	11	14.6	31.6	11	10	1	0	0	17	15	7	7	10	21

Table S5.1. (continued).

sp#	Family	Species	Cluster Group	IndVal	Area (in %)	1 14.2	2 4.4	3 5.2	4 5.7	5 8.4	6 8.9	7 6.8	8 14.2	9 7.8	10 12.8	11 11.7
1069	Meliaceae	*Aglaia spectabilis* Miq. S.S.Jain & Bennet	11	14.2	15.76	2	0	0	9	0	0	4	1	24	29	31
265	Thymelaeaceae	*Enkleia malaccensis* Griff.	11	14.1	8.91	7	0	0	0	0	1	1	3	22	24	42
696	Anacardiaceae	*Pegia sarmentosa* Lecomte Hand.-Mazz.	11	14.1	12.46	27	0	0	0	0	1	9	0	5	22	37
184	Euphorbiaceae	*Spathiostemon javensis* Blume	11	13.9	21.42	15	0	0	0	6	0	2	7	20	22	27
917	Magnoliaceae	*Magnolia dolichogyna* Dandy ex Noot. Figlar & Noot.	11	13.6	12.18	35	0	0	0	0	0	0	11	0	15	38
943	Polygalaceae	*Xanthophyllum impressum* Meijden	11	13.6	8.24	24	0	0	0	0	0	0	1	2	28	45
1199	Myristicaceae	*Horsfieldia polyspherula* Hook.f. emend. King J.Sinclair	11	13.4	4.96	2	0	0	0	4	0	9	1	7	22	55
830	Dipterocarpaceae	*Shorea retusa* Meijer	11	13.3	9.25	0	1	0	0	0	0	14	6	16	22	39
1026	Sapindaceae	*Lepisanthes amoena* Hassk. Leenh.	11	12.7	49.89	3	11	1	0	3	13	10	12	16	15	16
1251	Moraceae	*Artocarpus teysmannii* Miq.	11	12.4	17.84	0	0	0	1	25	1	10	5	24	9	26
1028	Sapindaceae	*Lepisanthes kinabaluensis* Leenh.	11	12.1	13.3	29	0	1	0	9	0	4	10	0	14	33
852	Dipterocarpaceae	*Vatica javanica* Slooten subsp. *scaphifolia* Kosterm. P.S.Ashton	11	11.6	7.71	0	0	0	0	29	0	0	1	33	0	38
344	Dichapetalaceae	*Dichapetalum gelonioides* Roxb. Engl. subsp. *tubetculatum* Leenh.	11	11	20.43	23	0	1	0	0	0	2	11	19	19	25
390	Proteaceae	*Helicia maxwelliana* Gibbs	11	10.4	11.59	32	0	1	0	0	0	0	7	8	18	33
886	Liliaceae	*Smilax corbularia* Kunth var. *woodii* Merr. T.Koyama	11	10	7.04	37	0	0	0	0	0	0	1	0	19	43
538	Oxalidaceae	*Sarcotheca rubrinervis* Hallier f.	11	9.8	3.68	8	0	0	0	8	0	0	2	0	26	57
567	Fagaceae	*Lithocarpus grandifolius* D.Don S.N.Biswas	11	9.7	11.87	15	0	0	0	0	0	0	4	28	23	30
552	Fagaceae	*Castanopsis megacarpa* Gamble	11	9.5	10.26	26	0	3	0	0	5	2	1	0	30	34
1105	Caesalpiniaceae	*Afzelia rhomboidea* Blanco Vidal	11	8.8	7.18	21	0	0	0	0	0	0	2	0	37	40
542	Violaceae	*Rinorea lanceolata* Wall. Kuntze	11	8.5	13.47	21	0	0	0	17	0	0	9	0	25	28
662	Ulmaceae	*Celtis philippensis* Blanco	11	8.4	7.5	8	0	0	0	0	0	0	18	0	36	38
225	Dilleniaceae	*Saurauia agamae* Merr.	11	6.2	9.46	24	0	0	0	0	0	0	11	12	24	29
795	Dipterocarpaceae	*Shorea gratissima* Wall. ex Kurz Dyer	11	6.2	3.22	1	0	0	0	0	0	0	1	8	42	48
193	Caprifoliaceae	*Viburnum amplificatum* Kern	11	5.1	2.67	17	0	0	0	0	0	0	0	0	34	49
670	Anacardiaceae	*Anacardium occidentale* L.	11	3.3	1.91	0	0	0	0	18	0	0	0	10	28	44

Samenvatting

De biodiversiteit van Borneo, het derde grootste eiland ter wereld, wordt ernstig bedreigd door houtkap en veranderd landgebruik. Ondanks de erkenning van Borneo als een van de belangrijkste centra van biodiversiteit, is er betrekkelijk weinig bekend over de ruimtelijke patronen van a) plantendiversiteit en b) de verschillende floristische regio's. Wat er wel bekend is, is bovendien hoofdzakelijk gebaseerd op 'subjectieve' specialistische kennis in plaats van op feitelijke data. Onlangs heeft het Nationaal Herbarium Nederland zijn plantencollecties van Borneo gedigitaliseerd. Deze dataset maakt het mogelijk om de ruimtelijke patronen op basis van werkelijke kwantitatieve data te onderzoeken – het onderwerp van dit proefschrift.

Voorwaarden voor analyse

Voordat gedigitaliseerde collecties gebruikt kunnen worden voor analyses van ruimtelijke patronen van diversiteit moeten ze aan twee voorwaarden voldoen:

1. Ze moeten correct geïdentificeerd zijn.
2. Ze moeten voorzien zijn van nauwkeurige geografische coördinaten.

Om aan de eerste voorwaarde te voldoen, hebben we onze analyses beperkt tot soorten die behoren tot plantenfamilies die in 'Flora Malesiana' bewerkt zijn, met andere woorden soorten die door een specialist voorzien zijn van een correcte soortnaam. Hieraan toegevoegd zijn de collecties behorend tot de bewerkte geslachten van de zuurzakfamilie (Annonaceae), de wolfsmelkfamilie (Euphorbiaceae), en de orchideeën (Orchidaceae).

Om aan de tweede voorwaarde te voldoen, hebben we zoveel mogelijk vindplaatsen opgezocht in zogenaamde gazetteers. Dit zijn lijsten met plaatsnamen en corresponderende coördinaten. Na plaatsing van alle vindplaatsen op de kaart van Borneo bleek dat een onevenredig aantal collecties uit noordelijk en westelijk Borneo kwam (figuur 3.3). Om deze onevenredigheid tot een minimum te beperken, hebben we zoveel mogelijk vindplaatsen uit zuidelijk en oostelijk Borneo (die nog niet in gazetteers vertegenwoordigd zijn) voorzien van coördinaten. Hiervoor gebruikten we de techniek 'georegistratie'.

Georegistratie

De collecties uit zuidelijk en oostelijk Borneo zijn veelal gemaakt rond het einde van de 19e eeuw en het begin van de 20e eeuw. Tegenwoordig hebben veel van deze vindplaatsen een andere naam of bestaan ze niet meer. Gelukkig is er ook een groot aantal expeditiekaarten uit die tijd bewaard gebleven waar deze locaties op afgebeeld zijn. Door deze (gedigitaliseerde) expeditiekaarten te koppelen aan satellietkaarten, konden we de vindplaatsen van zeer nauwkeurige coördinaten voorzien. Dit heeft ertoe geleid dat we 2.577 extra collecties, behorend tot 1.744 soorten en verzameld op 134 'nieuwe' locaties, konden toevoegen aan de dataset. Dit heeft geresulteerd in een totale dataset van 66.262 bruikbare collecties, behorend tot 102 verschillende plantenfamilies. Dit proces wordt in detail beschreven in hoofdstuk 2.

Soort-verspreiding-modellen

Om ruimtelijke patronen van plantendiversiteit en -zeldzaamheid vast te stellen, hebben we Borneo allereerst verdeeld in rastercellen van circa 100 km2 . Hoewel alles in het werk is gesteld om zoveel mogelijk vindplaatsen van coördinaten te voorzien, zijn in 'slechts' 1.837 van de totaal 8.577 cellen van Borneo (21.4%; figuur 3.3) collecties aangetroffen die vertegenwoordigd zijn in de dataset van het Nationaal Herbarium Nederland. Om de aan- of afwezigheid van individuele soorten toch te kunnen vaststellen – ook voor die cellen waarvoor geen collecties beschikbaar zijn – hebben we zogenaamde 'soort-verspreiding-modellen' gebruikt. Het bijzondere aan deze modellen is dat zij uitsluitend data gebruiken over de aanwezigheid van plantensoorten en dat gegevens over de afwezigheid van die soorten (wat zelden met zekerheid vast te stellen is) niet nodig zijn.

Soort-verspreiding-modellen identificeren relaties tussen omgevingsomstandigheden (zoals de gemiddelde temperatuur en de totale hoeveelheid neerslag) op locaties waar een plantensoort daadwerkelijk gevonden is. Door extrapolatie van deze relaties voorspellen zij de potentiële aan- en afwezigheid van deze soort voor geheel Borneo. Voorbeeld: de locaties waar een plantensoort gevonden is kenmerken zich door een hoogte tussen 1.200-1.750 meter en een jaarlijkse neerslag tussen 1.500-2.000 mm. Door de gebieden op Borneo te identificeren die aan deze voorwaarden voldoen, wordt het mogelijk de potentiële verspreiding van deze soort te voorspellen.

Null-model

Problematischer was het om de soort-verspreiding-modellen statistisch te testen. De bestaande tests gaan allen uit van data over zowel de aan- als afwezigheid van soorten. In ons geval ontbraken de laatste gegevens, waardoor deze tests feitelijk onbruikbaar waren. Vaak wordt dit opgelost door 'schijnafwezigheid'-data te genereren. Dit zijn willekeurig geselecteerde locaties waar geen aanwezigheid is geconstateerd voor de soort in kwestie, zodat de statistische tests toch bruikbaar zijn. Na studie bleek echter dat de standaard statistische waarden, die gebaseerd zijn op aanwezigheid- en schijnafwezigheid-data, onder deze omstandigheden niet bruikbaar waren. Om deze modellen toch op statistische significantie te kunnen testen, hebben we een null-model methode geïntroduceerd (zie hoofdstuk 3). Deze methode test of door toeval een model ontstaan kan zijn dat even goed is als het soort-verspreiding-model van een werkelijke soort.

Stel: een plantensoort komt in de dataset op 12 locaties voor. Op basis hiervan wordt een soort-verspreiding-model gemaakt. Van dit model wordt de statistische waarde bepaald. Vervolgens worden 1.000 keer 12 willekeurige locaties van Borneo getrokken. Voor ieder van deze 1.000 sets van 12 willekeurige locaties wordt een soort-verspreiding-model ontwikkeld op dezelfde manier als voor de werkelijke soort is gedaan. Voor alle 1.000 modellen worden ook de statistische waarden bepaald en deze worden gerangschikt van klein naar groot. Als de statistische waarde van het werkelijke soort-verspreiding-model groter is dan de 950e waarde van de 1.000 modellen gebaseerd op willekeurig getrokken locaties, dan is de kans dat door toeval een zelfde resultaat zou kunnen zijn ontstaan kleiner dan 5% (de standaard statistische ondergrens).

Alle soort-verspreiding-modellen zijn op deze wijze getest. Dus een soort vertegenwoordigd door 13 aanwezigheid locaties is getest tegen de 950e waarde van 1.000 keer 13 willekeurige locaties, enzovoorts. Dit houdt in dat voor een significant soort-verspreiding-model een relatie met zijn omgevingsomstandigheden geïdentificeerd wordt die niet door toevallige omstandigheden verklaard kan worden.

Patronen van plantendiversiteit en -zeldzaamheid
In hoofdstuk 4 wordt beschreven hoe we patronen van plantendiversiteit en -zeldzaamheid hebben ontwikkeld op basis van alle significante soort-verspreiding-modellen. Om relaties tussen het voorkomen van een soort en zijn omgevingsomstandigheden te kunnen vaststellen, is een minimum aantal collecties noodzakelijk. Wij hebben ons beperkt tot soorten die zijn vertegenwoordigd door minstens vijf unieke aanwezigheid locaties (een soort is maximaal één keer vertegenwoordigd in iedere cel). Aan deze voorwaarde voldeden 2.273 soorten, vertegenwoordigd door 44.106 collecties, variërend van 5-202 unieke locaties per soort. Van de 2.273 soort-verspreiding-modellen waren er 1.439 significant (63.3%). Het plantendiversiteitpatroon is ontwikkeld door alle 1.439 modellen over elkaar te leggen (figuur 4.2A).

Het zeldzaamheidpatroon is ontwikkeld door de inverse van de soort-verspreiding-modellen over elkaar te leggen (figuur 4.2B). Een zeldzame soort die op basis van zijn soort-verspreiding-model in 100 cellen op Borneo aanwezig is, krijgt voor ieder van deze cellen een waarde van 1/100. Een algemenere soort die in 1.000 cellen aanwezig is, krijgt voor iedere cel een waarde van 1/1.000 waardoor zijn aandeel in het patroon van zeldzaamheid veel kleiner is. Het kan worden verwacht dat op locaties waar veel soorten voorkomen ook veel zeldzame soorten vertegenwoordigd zijn. Door zeldzaamheid tegen diversiteit uit te zetten in een grafiek (figuur 4.4) is het mogelijk de hotspots te identificeren: gebieden waar meer zeldzame planten voorkomen dan op basis van de diversiteit verwacht kan worden.

Daarnaast hebben we onderzocht welke omgevingsfactoren kenmerkend zijn voor:
a) gebieden met een hoge plantendiversiteit;
b) gebieden waar veel zeldzame planten voorkomen;
c) de hotspots.
Gebieden met een hoge diversiteit en gebieden waar veel zeldzame planten voorkomen worden gekenmerkt door kleine temperatuurverschillen gedurende het jaar. Binnen dit beperkte bereik zijn er echter wel seizoensverschillen. De hotspots worden gekenmerkt door hoogte (bergachtigheid), de structuur van de bodem en de koolstof-stikstofverhouding van de bodem.

Floristische regio's
In hoofdstuk 5 beschrijven we hoe we de floristische regio's van Borneo hebben geïdentificeerd (gebieden met een relatief uniforme soortensamenstelling). We hebben dit gedaan aan de hand van een clusteranalyse gebaseerd op de 1.439 significante soort-verspreiding-modellen. Voor ieder van de 8.577 cellen van Borneo weten we welke van de 1.439 soorten voorspeld worden voor te komen. Op basis van deze gegevens konden we voor iedere combinatie van twee cellen berekenen in hoeverre ze in soortensamenstelling overeenkomen. De maat van overeenkomst die we gebruikt hebben is de 'Sørensen index'. Een clusteranalyse plaatst cellen die veel overeenkomstige soorten bevatten in één groep. Het resultaat is een clusterdiagram dat veel lijkt op een boom met takken. Het eerste clusterniveau verdeelt de 8.577 cellen in twee groepen op basis van de Sørensen indices. Vervolgens splitsen beide takken wederom in twee groepen, enzovoorts, net zolang tot dat de 8.577 cellen niet verder verdeeld kunnen worden. De vraag is nu: op welk niveau van de clusterboom zijn de clustergroepen het meest representatief voor de floristische regio's?

Om die vraag te beantwoorden, hebben we de 'indicator-soort-analyse' toegepast. Deze analyse maakt gebruik van de indicatorwaarde van iedere soort. De indicatorwaarde is maximaal wanneer een soort uitsluitend in één clustergroep en tevens in alle cellen van deze groep voorkomt. In het geval van twee clustergroepen zullen veel soorten uitsluitend in één van de twee clustergroepen voorkomen, maar waarschijnlijk niet in alle cellen behorend tot die groep. In het andere uiterste (veel clustergroepen) zal het voorkomen van veel soorten niet beperkt zijn tot één clustergroep, maar zullen deze wel in alle cellen van een aantal groepen voorkomen. Door de indicator-soort-analyse op ieder clusterniveau uit te voeren (hier 2-40 clustergroepen) konden we vaststellen op welk clusterniveau gelijktijdig voor alle soorten optimaal aan beide criteria wordt voldaan. Voor het geval van Borneo was dit zo bij 11 clustergroepen (figuur 5.1). De resulterende clusterboom en de geografische positie van de corresponderende rastercellen zijn afgebeeld in figuur 5.2.

De belangrijkste verschillen met de huidige indeling van de floristische regio's door experts (figuur S5.2 en kaft achterzijde) is dat het laagland regenbos opgedeeld moet worden in vier (en mogelijk zes) verschillende regio's. De condities van de omgevingsfactoren waaronder de verschillende floristische regio's voorkomen zijn geïdentificeerd met een 'CART analyse' (figuur 5.3).

Conclusies en aanbevelingen
In hoofdstuk 6 vatten we alle resultaten van het onderzoek samen en analyseren we welke percentages van gebieden met hoge plantendiversiteit en -zeldzaamheid, de hotspots en de verschillende floristische regio's tegenwoordig verdwenen zijn door houtkap en/of veranderd landgebruik. Ook geven we aan welk percentage van deze gebieden beschermd is in natuurgebieden die door IUCN (International Union for Conservation of Nature) erkend zijn.

Om dit te kunnen doen, hebben we allereerst de patronen van diversiteit en zeldzaamheid opgedeeld in vier gelijke kwartielen (0-25, 25-50, 50-75, 75-100%, figuur 6.1A&C). Voor de schatting van natuurlijk bos op Borneo hebben we de dataset 'Global Land Cover 2000' gebruikt. Dit is een 1 km resolutie kaart met een classificatie van bostypen (variërend van natuurlijke tot ernstig verstoorde of helemaal verdwenen vegetatie) die is afgeleid van satellietgegevens van de jaren 1998-2000. Deze schatting is dus conservatief. Op basis van deze gegevens is de conclusie dat reeds 36.4% van Borneo is ontbost. Ongeveer 10% van het oppervlak van Borneo heeft een beschermde status (tabel 6.1). Hiervan is echter al 18.4% verdwenen. Dramatischer is dat 57.4% van de meest diverse gebieden reeds ontbost is, terwijl dit gebied maar 5.6% van Borneo's oppervlak beslaat. Van deze meest diverse gebieden heeft 11% een beschermde status. Zelfs van dit gebied is reeds één derde verdwenen. Vergelijkbare getallen zijn aangetoond voor de gebieden die de meeste zeldzame soorten herbergen. Voor de hotspots zijn de getallen minder ernstig, omdat die gebieden zich hoofdzakelijk in de moeilijk toegankelijke centrale bergketen bevinden, de regio die het minst getroffen is door ontbossing. De floristische regio's 1, 2 en 3, die zich in het gebied van de centrale bergketen bevinden, herbergen dan ook nog meer dan 90% van hun natuurlijke vegetatie. Daarentegen hebben bijna alle laagland floristische regio's (met uitzondering van regio 6) meer dan 45% van hun vegetatie reeds verloren. Zelfs hun beschermde gebieden zijn ernstig aangetast.

Met dit onderzoek naar de patronen van plantendiversiteit en -zeldzaamheid, hotspots en floristische regio's van Borneo gebaseerd op data van het Nationaal Herbarium Nederland, verschaffen wij natuurbeschermingsorganisaties en regeringen solide gegevens om belangrijke delen van Borneo's natuurlijke botanische rijkdom te behouden. Onze aanbeveling is om de meest diverse gebieden, de hotspots en gebieden met het hoogste percentage aan zeldzame soorten een beschermde status te verlenen. Daarnaast zou een aanzienlijk deel van ieder van de 11 floristische regio's een beschermde status moeten krijgen, die dan vervolgens ook gehandhaafd dient te worden.

Curriculum Vitae

Niels Raes was born on the 9th of March 1971 in Bladel, the Netherlands. At the age of 1 year he moved with his parents to Hapert, where he attended primary school. At an age of 12 years he went to the Pius X high school in Bladel where he concluded the VWO exam in 1990. The same year he started his study Biology at Utrecht University. In 1995 he had his first opportunity to do fieldwork in the tropics. Together with Menno de Lind van Wijngaarden, and under supervision of Dr. Renske Ek, he studied the impact of selective logging and beta diversity of lianas (climbing plants) in the tropical rain forests of Guyana, South America. In 1998 he obtained his MSc. degree in Tropical Ecology and Tropical Botany.

After his studies, Niels was first a computer teacher for adults at 'Bamboe Opleidingen' in Utrecht. It was here that he developed his computer skills, which proved beneficial for his PhD research later on. He left this job for one month, after an invitation of Dr. Bruno Bordenave to join a botanical exploration of the Kaw Mountains in French Guiana. The purpose of this exploration was to identify high conservation value areas within the ASARCO mining company concession. In these mountains they discovered a new plant genus named *Hekkingia* (species *bordenavei*) belonging to the Violaceae plant family.

In 2000 he was given the opportunity to continue his biology career at Utrecht University by participating in the EU funded liana research of Dr. Renske Ek. He returned to Guyana to continue the field research on lianas, this time supervising two MSc. students. After having successfully finished this project for Utrecht University he was unemployed for 9 months. During this period he developed, together with Dr. Renske Ek, the 'Climbers of Guyana'-website that contains the botanical descriptions and pictures of most lianas that occur in central Guyana.
At the end of 2002 a research proposal to study the impact of selective logging on lianas in Bolivia was granted by USAID. The next 9 months he conducted fieldwork in Bolivia under supervision of BOLFOR (Bolivia Forestal).

Then, in November 2003 he started his PhD project entitled 'Plant diversity of Borneo deduced from collection databases' at the National Herbarium of the Netherlands, Leiden University, which led to this thesis. Besides acquiring the statistical modelling skills leading to the patterns of biodiversity presented in this thesis, he also joined a botanical expedition to the 'Gunung Lumut Protection Forest'. This gave him the opportunity to see the real, beautiful, and threatened tropical rain forests of Borneo.

Publications

Raes, N., van Loon, E.E., Roos, M.C. and ter Steege, H. (subm.). The floristic regions of Borneo inferred from species distribution models.

Slik, J.W.F., Aiba, S.-I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K., Nagamasu, H., Payne, J., Paoli, G., Poulsen, A., **Raes, N.**, Sheil, D., Suzuki, E. (subm.). Forest Carbon and Biogeography: Explaining why Bornean Forests Possess 60% More Live Aboveground Biomass than the Amazon.

Slik, J.W.F., Aiba, S.-I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K., Nagamasu, H., Nilus, R., Payne, J., Paoli, G., Poulsen, A., **Raes, N.**, Sheil, D., Sidiyasa, K., Suzuki, E., van Valkenburg, J.L.C.H., Webb, C.O., Wilkie, P. and Wulffraat, S. (subm.). Wood specific gravity, stem density, basal area and aboveground biomass gradients and their environmental correlates in Borneos tropical forests.

Raes, N., Roos, M.C., Slik, J.W.F., van Loon, E.E. and ter Steege, H. (accepted). Botanical richness and endemicity patterns of Borneo derived from species distribution models. Ecography.

Raes, N., Mols, J.B., Willemse, L.P.M. and Smets, E.F. (accepted). Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images. Blumea 54(1).

Raes, N. and van Welzen, P.C. (accepted). The demarcation and internal division of Flora Malesiana: 1857- Present. Blumea 54(1).

Slik, J.W.F., **Raes, N.**, Aiba, S.-I., Brearley, F.Q., Cannon, C.H., Meijaard, E., Nagamasu, H., Nilus, R., Paoli, G., Poulsen, A.D., Sheil, D., Suzuki, E. , van Valkenburg, J.L.C.H., Webb, C.O., Wilkie, P., and Wulffraat, S. (accepted). Environmental correlates for tropical tree diversity and distribution patterns in Borneo. Diversity and Distributions.

Raes, N. and ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. Ecography 30(5): 727-736.

Slik, J.W.F., Hovenkamp, P., Iqbal, M. and **Raes, N.** (2007). Structure, plant species diversity and plant species composition of the Gunung Lumut Protection Forest. In: Options for biodiversity conservation and sustainable use in lowland forest of southeast Borneo; Proceedings of a workshop organized on 19 May 2006 in Leiden, the Netherlands (Eds. Hans H. de Iongh, Gerard A. Persoon and Wawan Kustiawan), pp. 71-107. ISBN 978 90 5191 153 4.

Zagt, R.J., Ek, R.C. & **Raes, N.** (2003). Logging and Lianas – Tropenbos-Guyana Series 2003-1. Tropenbos International, Wageningen, The Netherlands.

Raes, N. & Ek, R.C. (2001). Climbers of Guyana – website. http://www.bio.uu.nl/~herba/lianas

Bordenave, B., **Raes, N.** & de Granville, J.-J. (2000). Etat Initial de la Vegetation Forestière de la Montagne de Kaw – 2. Convention IRD/ASARCO n° 9051.

Acknowledgements

My interest for the tropics - with its exotic animals, like monkeys and colourful birds, the enormous trees, beautiful flowers, and extended coral reefs, but also with different human cultures - was raised at an early age. In 1995, the moment was finally there that I, together with Menno de Lind van Wijngaarden, was given the chance to do research in the tropical forests of Guyana. Our supervisor at Utrecht University, Renske Ek, gave us the opportunity to study the impact of selective logging on liana diversity in the tropical forests of Guyana. Working with local field assistants, (amongst others) Milton 'Suki' Allicock and Martin 'short man' Williams (who sadly passed away in 2001), gave me understanding of forest ecology not described in text books. While being friends during leisure time, they also provided me a glimpse on the Guyanese Amerindian culture. I would like to thank them, and the staff of the NHN-Utrecht (Paul and Hiltje Maas, Marion Jansen-Jacobs, Lubbert Westra, Bep Mennega, Jifke Koek-Noorman, and all others) for the support and the given opportunity to start a career as tropical botanical ecologist.

After having finished my MSc., I was able to continue to work as tropical ecologist by joining a botanical inventory of the Kaw Mountains in French Guiana led by Bruno Bordenave (2000), by returning to Guyana to continue the liana work of Renske Ek (2001), and by conducting liana research for BOLFOR in Bolivia supervised by Marielos Peña-Claros and Francis 'Jack' Putz (2003). Thank you for having provided these opportunities and support to continue my career. Not to forget Olaf Bánki, Tinde van Andel, and Renske Ek with whom we almost erected the 'Hubudi' foundation with the purpose to carry out consultancy biodiversity assessments, keeping my spirit up in times I was not employed as an ecologist.

Then in November 2003 I started my PhD at the National Herbarium of the Netherlands – Leiden University branch. Although sometimes difficult, I do appreciate the freedom given by my promotor Pieter Baas, and co-promotores Marco Roos en Hans ter Steege, to examine different strategies to tackle the problem of developing the high spatial resolution diversity patterns of Borneo which has finally resulted in the methodological approaches as I present them in this thesis. Notwithstanding that they were always available to give their advise and opinions on my lines of thought. Thank you for the advice, support, and trust.
Invaluable in the process leading to this thesis were also Emiel van Loon, Ferry Slik, Peter van Welzen, Peter Hovenkamp, Johan Mols, Soraya Sierra, Max van Balgooy, Paul Keßler, and the colleagues I met at (inter)national conferences, of which I particularly like to thank Joaquín Hortal and Jorge Lobo for their discussions on null-models. Furthermore, I would like to thank Bart Meganck, Chuck Cannon, and Gary Paoli (and the anonymous reviewers) for the suggested improvements and critical comments to earlier versions of the chapter manuscripts. Special mention here also deserve Ben Kieft, who did an excellent job in digitising the old expedition maps used for my research, Jaap Vermeulen who kindly provided his Orchid database, and the Manifold GIS Community who helped me out on numerous GIS problems.
And there were also the students I supervised, Muhammed Iqbal, Rebecca Hsu, and Jan van Voorstvader, who helped me with several problems and furthered my insights in the possibilities and drawbacks of using species distribution models.
Writing a thesis on the biodiversity of Borneo would not have been complete without visiting this remarkable island. Tropenbos Indonesia provided me the opportunity to participate in a large biodiversity assessment of the Gunung Lumut Protection Forest in 2005, organized by Hans de Iongh. Working with very enthusiastic and skilful local guides (and tree-climbers) like Pa Didan, Indonesian botanical experts Kade Sidiyasa, Arbainsyah, Ambriansyah, and Zainal Arifin and the colleagues of Naturalis and the NHN, Peter Hovenkamp, Ferry Slik, Jan van Tol, Erik van Nieukerken, and Rienk de Jong was unforgettable.

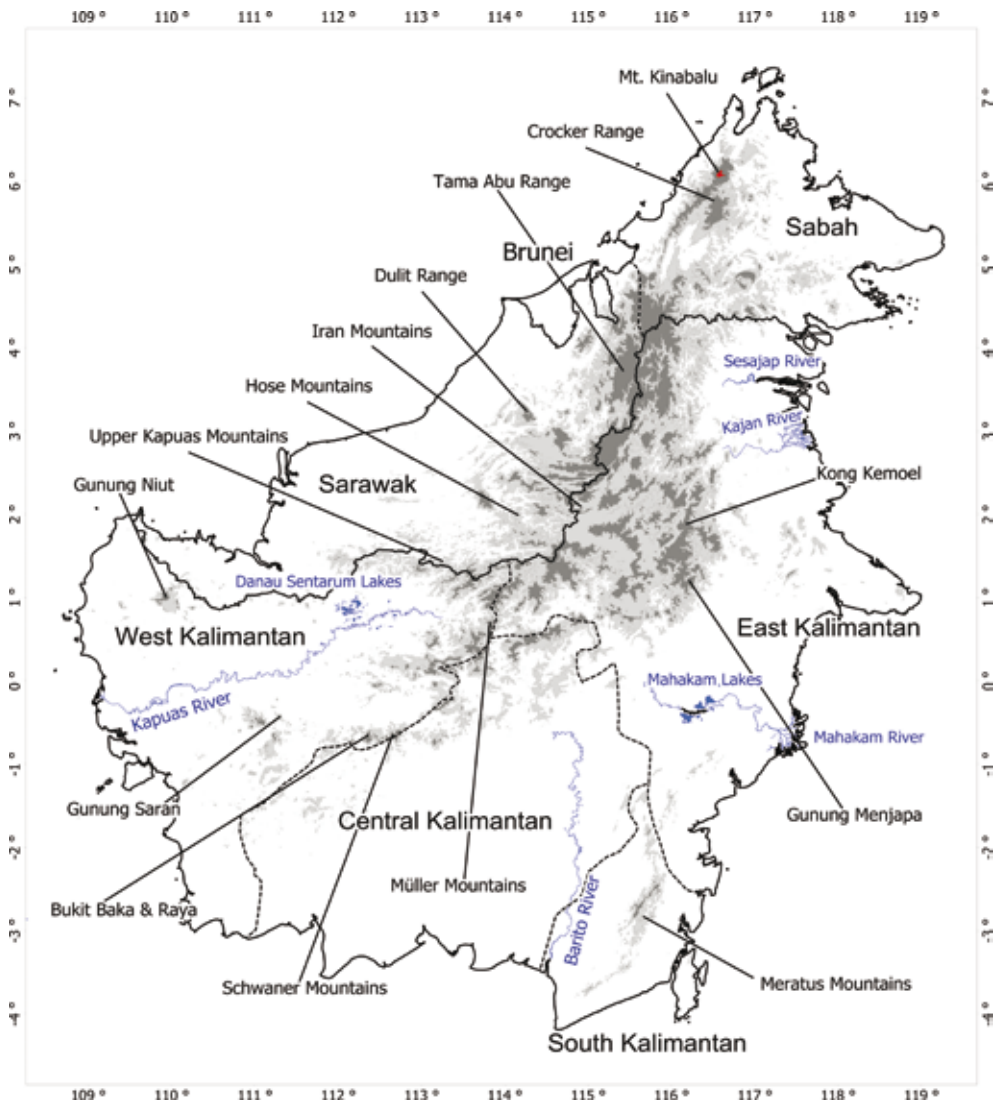
Without the professional editing of the Dutch summary by Koen-Machiel van de Wetering it would not have been as comprehensive as it is published in my thesis. Even more credits go to Ed van Oosterhout who is responsible for the graphic design of my thesis. Then, there were always my friends who are almost experts on biodiversity issues by now, I guess. Thank you for your interest and patience to listen to my stories and frustrations.

Mijn ouders kunnen hier ook niet ontbreken. Hoewel jullie bijdrage hoofdzakelijk indirect is geweest (afgezien van de wiskundebijles op het gebied van 'species-area relationships'), is deze daarom zeker niet minder belangrijk. Jullie hebben mij, mijn broer Coen en zus Susan grootgebracht met een liefde voor de natuur, reizen, en cultuur. Jaarlijks gingen we driemaal op vakantie, waarvan in de zomer zeker 5 á 6 weken. Een duidelijke bestemming was er nooit, en na 3 á 4 dagen ergens te zijn geweest ging onze 'ontdekkingsreis' weer verder. Zo hebben wij grote delen van Europa al op jonge leeftijd gezien. Wandeltochten gingen bij voorkeur buiten de gebaande paden, en toch kwamen we altijd weer uit bij het begin. Daarnaast heb ik mooie herinneringen aan de vistochten met mijn vader, waarbij het de sport was om de meest afgelegen stukjes rivier op te zoeken om daar te gaan vissen. Deze ervaringen hebben mij in belangrijke mate gevormd. Ook later zijn jullie mij op allerlei mogelijke manieren blijven ondersteunen. Hierdoor kon ik vol vertrouwen elke keer weer nieuwe uitdagingen aangaan, waaronder de meest recente - dit proefschrift. Ik wil jullie daar heel hartelijk voor bedanken.

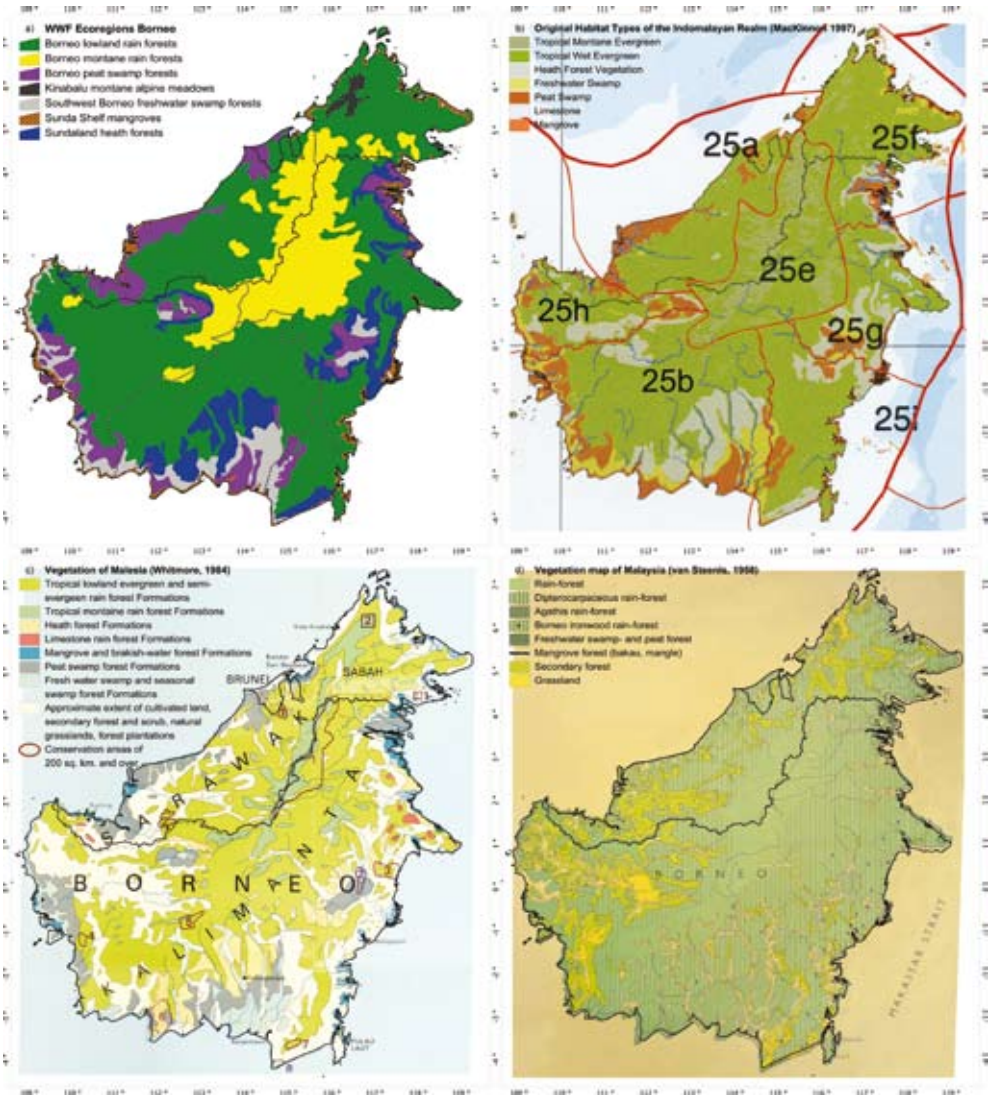
Then finally, I want to thank you – Linda – my love. You have supported and motivated me to finish my thesis. You had to withstand my grumpy moods when writing would not progress as fast as I had in mind, or when I received reviewer comments with the result that I had to rewrite part of the thesis and resulting papers (in the end improving the content). At important moments you showed genuine interest, or came up with welcome distractions and holidays which allowed me to clear my head. Thank you, and the many people who have contributed to my thesis but are not mentioned personally here.

Niels

Koompassia excelsa Taub. Picture taken in the Gunung Lumut Protection Forest East Kalimantan, Indonesia.



Map indicating the Bornean administrative provinces of Malaysia and Indonesia, the country of Brunei, the major mountain ranges and large rivers. Areas between 500-1000m in light-grey; areas above 1000m in dark-grey; Mt. Kinabalu in red (≥ 2600m) (derived from the SRTM 90m resolution altitude data (<http://srtm.csi.cgiar.org>)).



A) WWF ecoregions of Borneo; colours similar to floristic regions map (Chapter 5).

B) Detail from the 'Original Habitat Types of the Indomalayan Realm' of MacKinnon (1997).

C) Detail from the 'Vegetation of Malesia' map of Whitmore (1984).

D) Detail from the 'Vegetation map of Malaysia' of van Steenis (1958).

All maps in geographic projection.