Bohlman, S.A., W.F. Laurance, S.G. Laurance, H.E.M. Nascimento, P.M. Fearnside & A. Andrade. nd. Effects of soils, topography and geographic distance in structuring central Amazonian tree communities. *Journal of Vegetation Science* (in press).

# Effects of soils, topography and geographic distance in structuring central Amazonian tree communities

**Stephanie A. Bohlman**[1,2*]; **William F. Laurance**[1,3,4]; **Susan G. Laurance**[1,5]; **Henrique E. M. Nascimento**[3,6]; **Philip M. Fearnside**[7,8] & **Ana Andrade**[3,9]

[1]*Smithsonian Tropical Research Institute, Apartado Postal 2072, Balboa, Panama;*
[2]*Present address Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544 USA;* [3]*Biological Dynamics of Forest Fragments Project, National Institute for Amazon Research (INPA), C.P. 478, 69011-970 Manaus, AM, Brazil;* [5]*LauranceS@si.edu;* [6]*henrique@inpa.gov.br;*
[7]*Department of Ecology, National Institute for Amazon Research (INPA), C.P. 478, 69011-970, Manaus, AM, Brazil;* [8]*pmfearn@inpa.gov.br;* [9]*titina@inpa.gov.br;*
[*]*Corresponding authors; E-mail: sbohlman@princeton.edu*

## Abstract

**Question:** What is the relative contribution of geographic distance, soil and topographic variables in determining the community floristic patterns and individual tree species abundances in the nutrient-poor soils of central Amazonia?

**Location:** Central Amazonia near Manaus, Brazil

**Methods:** Our analysis was based on data for 1105 tree species (≥10 cm dbh) within 40 1-ha plots over a ~1000-km$^2$ area. Slope and 26 soil-surface parameters were measured for each plot. A main soil-fertility gradient (encompassing soil texture, cation content, nitrogen and carbon) and five other uncorrelated soil and topographic variables were used as potential predictors of plant-community composition. Mantel tests and multiple regressions on distance matrices were used to detect relationships at the community level, and ordinary least square (OLS) and conditional autoregressive (CAR) models were used to detect relationships for individual species abundances.

**Results:** Floristic similarity declined rapidly with distance over small spatial scales (0-5 km), but remained constant (~44%) over distances of 5 to 30 km, which indicates lower beta diversity than in western Amazonian forests. Distance explained 1/3 to 1/2 more variance in floristics measures than environmental variables. Community composition was most strongly related to the main soil-fertility gradient and C:N ratio. The main fertility gradient and pH had the greatest impact of species abundances. About 30% of individual tree species were significantly related to one or more soil/topographic parameters.

**Conclusions:** Geographic distance and the main fertility gradient are the best predictors of community floristic composition, but other soil variables, particularly C:N ratio, pH, and slope, have strong relationships with a significant portion of the tree community.

**Keywords**: Beta diversity; Brazil; floristics; tropical forest; species abundances; spatial dependence

# Introduction

Much research has focused on the relative role of distance-related processes such as dispersal versus environmental factors in shaping community floristics in tropical forests (Clark et al. 1999; Chust et al. 2006; Hubbell 2001; Phillips et al. 2003; Potts et al. 2002; Webb & Peart 2000 and others). Certainly no one answer has emerged as to which is more important. Instead, the importance of different factors varies from region to region (Palmer 2006; Tuomisto et al. 2003a, b, c), and even at different spatial scales in the same region (Normand et al. 2006; Potts et al. 2002), in response to past and recent geologic history, human and natural disturbance patterns, and evolutionary processes.

Nonetheless, some environmental variables, particularly soil texture, cation composition and topography, are repeatedly reported as important in determining tropical forest composition (Costa et al. 2005; Sri-Ngernyuang et al. 2003; Tuomisto et al. 2002; Valencia et al. 2004; Vormisto et al. 2004). The availability of complete floristics and soils data sets that are needed to address these questions is very limited given the vast extent of the tropics. In the large area covering the central Amazon basin, only one study with detailed soil and floristic data examines the relationship between environmental variables and floristics (Costa et al. 2005) and this is limited to understory herbs. To put the floristic community in context of other studies in the Neotropics (Condit et al. 2000; Pitman et al. 2001; ter Steege et al. 2006), an analysis of trees (> 10 cm dbh) is needed.

In the relatively nutrient-rich soils of western Amazonia, evidence is accumulating that tree communities exhibit surprisingly little turnover in floristic composition (beta diversity) at larger spatial scales (10–1000 km), despite having high local species richness (alpha diversity) and relatively rapid turnover at local scales (<10 km) (Condit et al. 2002; Macia and Svenning 2005; Pitman et al. 2001). Species diversity patterns are less clear for the vast region of central Amazonia, where soil fertility tends to be very low (Brown 1987; Chauvel et al. 1987). Central Amazonia has one of the highest alpha diversities of trees recorded for any tropical forest (Oliveira & Mori 1999), but species turnover at landscape scales (distances of <100 km) has not yet been assessed.

Here, we describe the patterns of beta diversity in a central Amazonian forest and address two questions: What is the relative role of geographic distance and environmental variables in shaping floristic similarity and individual species abundances? Which soil and topographic variables are most important? Because spatial structuring in the environmental variables can cause inflated levels of environmental associations (e.g. Harms et al. 2001, Lichstein et al. 2002), our statistical approaches aimed to disentangle the effects of broad-scale geographic trends, local spatial autocorrelation and environmental variables. We also placed our findings in the context of other studies in the tropics. In particular, we examined how beta diversity patterns compare to other plot networks in the Neotropics, and how the proportion of species with significant soil or topographic relationships varied among different tropical forest communities.

## Methods

*Study area and floristics*

Our study area is part of the Biological Dynamics of Forest Fragments Project (BDFFP), a large-scale experimental study of habitat fragmentation, located in the central Amazon, 80 km N of Manaus, Brazil ($2^o$ 30' S, $60^o$ W) at 50-100 m elevation (Laurance et al. 2002; Lovejoy et al. 1986). The study area spans ~1000 km$^2$. The topography consists mostly of flat, clay-rich plateaus dissected by numerous stream and river gullies. Forests in the area are not seasonally inundated. Rainfall ranges from 1900-3500 mm annually with a pronounced dry season from June to October.

This site shows relatively few signs of human disturbance or large-scale natural disturbance in the past few hundred years. There is no evidence of swidden agriculture (Piperno & Becker 1996) and the majority of charcoal found at the site indicating fire dates from at least 1100–1500 yr ago (Piperno & Becker 1996; Santos et al. 2000). Large gaps created by wind bursts have been observed in this area of the central Amazon (Nelson et al. 1994) but there is no evidence of recent large blowdowns in the BDFFP plot network (Laurance et al. 2005; but see Nelson 2005).

Since 1980, a long-term study of tree community dynamics and composition has been conducted in fragmented and continuous forests in the study area. All trees $\geq$10 cm diameter-at-breast-height [dbh] are being monitored within 66 permanent, square, 1-ha plots spanning the experimental landscape. Live trees in each plot were marked and mapped and their diameters recorded, with measurements taken above the buttresses for buttressed trees. On average, 95.3% of all trees in each plot were identified to species (or morphospecies) level, using sterile or fertile material from each tree. Voucher specimens are housed in the BDFFP Plant Collection, Manaus, Brazil. The subset of 40 plots was used for this analysis contained 1105 identified tree species with a range of 224 to 293 species and 546 to 754 stems per plot.

Because we were interested in the relationship between floristics and environmental patterns in undisturbed forest, we used data only from the initial tree census, in the early-mid 1980s, completed before the BDFFP fragments were created and when the whole area was continuous forest. The distance between the 40 plots ranged from 0.10 to 32.5 km, with an average interplot distance of 13.8 km.

### *Soil data*

The dominant soils in the study area are xanthic ferralsols (using the FAO/UNESCO system; Beinroth 1975). Ferralsols are widespread in the Amazon Basin, heavily weathered, and usually have a low base saturation. They often are well aggregated, porous, and friable, with variable clay contents. Clay particles in ferralsols can form very durable aggregations, giving the soil poor water-holding characteristics, even with high clay contents (Richter & Babbar 1991). Xanthic ferralsols in the Manaus area are derived from Tertiary deposits and are typically acidic and very poor in nutrients such as P, Ca, and K (Chauvel et al. 1987).

A total of 21 soil parameters and 1 topographic parameter (slope) was collected from the plots. Each 1-ha plot was divided into 25 quadrats of 20 X 20 m each from which soil-surface samples (0-20 cm depth) were collected. The maximum slope within each quadrat was measured with a clinometer, then averaged to yield a mean value for the plot. To assess soil parameters, 9-13 quadrats were selected per plot, using an alternating pattern to provide good coverage of the plot. Within each quadrat, 15 surface samples were collected at haphazard locations using a soil auger, then bulked and subsampled. The organic matter layer was removed before sampling. The laboratory methods used for soil analyses are detailed in Fearnside and Leal-Filho (2001), and briefly summarized in Appendix 1.

Because many soil variables were intercorrelated, we used Principal Components Analysis (PCA) to identify major trends in the soil data using the 40 plots with complete soils data. The first PCA axis explained 50% of the total variance in the soil data and was strongly correlated with soil texture (fractions of clay, sand, and silt), total N, C content, total exchangeable bases, and several individual cation concentrations (Zn, Mg, Mn) (Appendix 2; Laurance et al. 1999). Because these parameters are strongly associated with soil fertility (Laurance et al. 1999), we hereafter refer to PCA axis 1, and the soil variables strongly correlated with it (particularly clay content), as the main soil-fertility gradient.

### *Statistical analyses*

In both the community and species level analyses, we sought to assess the relative importance of geographic distance (both broad- and fine-scale) and environmental variables in determining community composition and species abundance patterns. The community level analysis was based on interplot geographic distances and interplot differences in floristics and environmental variables. The individual species analysis was based on species abundances, geographic coordinates, and levels of the environmental variables in each plot. Interplot distances were used in the individual species analysis to measure fine-scale spatial autocorrelation among plots. We used one topographic and five soil variables in the analyses. One of the soil variables was clay content, which was employed as a surrogate for the main soil-fertility gradient because it was very strongly correlated with PCA axis 1 (r = -0.97) and was available for all 66 plots, rather than just the 40 plots used in the PCA. The other five variables, pH, C:N ratio, water content, phosphate, and slope, were only weakly correlated with the main soil-fertility gradient and with one another. Using Tukey's ladder of powers, the soil and topographic variables were transformed to reduce skewness (Table 1), so that the skewness of all transformed soil variables < |1|. Geographic distance was log-transformed to account for the main distance-dependent process, dispersal limitation, which is expected to cause logarithmic distance decay (Condit et al. 2002). Analyses were performed using the R 2.1.1 and S-Plus 7.0 statistical packages (www.R-project.org and Insightful Corporation, Seattle, WA, USA), including the S-Plus Spatial Module (Kaluzny et al. 1998).

### *Community level analyses*

We used three methods, based on interplot similarity matrices, to assess how changes in tree species composition are related to distance and soil composition/topography at the community level. First, we plotted interplot floristic similarity against interplot geographic distance and interplot similarity in soils. Second, we used Mantel tests to identify correlations between interplot similarity matrices (Legendre & Legendre 1998; Smouse et al. 1986). Mantel tests, which test for linear relationships between interplot differences, were deemed appropriate because plots of floristic similarity versus interplot differences in transformed distance and environmental variables showed no "hump-shaped" relationships. Partial Mantel tests were used to test the strength of relationship between floristic similarity and similarity in soil and topographic variables, while removing the portion of correlations based on interplot geographic distance alone. Third, we used multiple regressions on the interplot-distance matrices, described further below, to assess the relative importance of the soil and topographic variables and to partition the variance in floristic similarity among geographic distance and environmental factors (Legendre et al. 1994; Legendre & Legendre 1998).

Interplot floristic similarity was calculated with both Steinhaus' index, which incorporates species abundances, and Sørensen's index, which is based on presence/absence data (Legendre & Legendre 1998). For all three analyses, both indices yielded similar results. The results for Steinhaus' index are included in the text and those based on Sørensen's index can be found in the appendices. For the geographic-distance matrix, log-transformed Euclidean distances between plot-center coordinates were used. For the soil and topographic distance matrices, the differences between transformed values were used.

The change in floristic similarity with distance in the BDFFP plots was compared with distance decay in floristic similarity from plot networks in Panama, Ecuador and Peru (Condit et al. 2002; Pitman et al. 2001). The climate in Peruvian and central Amazonian sites average 2200-2300 mm/year with a 3 month dry season. The site in Ecuador is aseasonal with 3200 mm/yr rainfall on average. The Panama sites have the most geographic variation in rainfall (1800 – 3000 mm/rainfall). Methods in these other studies and this one were similar. In all plots, trees ≥ 10 cm dbh were used and all plots were 1 ha in size, except a single plot in Peru which is 0.875 ha. Because the species were identified by different botanists, there is potential for incongruent species identifications among the plot networks.

The relative contributions of distance and environmental variables were estimated using multiple regressions on the distance matrices. In this procedure, the distance matrix of interplot floristic similarity is the response variable, and the explanatory variables are interplot geographic distance and similarity in environmental variables. The matrix regression mimics normal regression, except the probability of the strength of relationship between response and explanatory variables is determined by permuting the original matrices 999 times and comparing the distribution of r values from the permutations against the observed value. Both forward and backward selection was used to choose the distance and environmental variables in the final model as follows. First, forward selection was used to add variables at a Bonferroni-corrected significant level of 0.005. Then backward elimination was applied to the model with all variables from the forward selection. The resulting final model was the most parsimonious model that explained the greatest variance in floristic similarity (RT) and included both distance and environmental and geographic variables together. We generated two additional models with a reduced number of variables. A distance-only model (RG) removed the environmental variables from the full model and the environment-only model (RE) removed distance. The results of these three models were used to partition the variation in floristic similarity into four fractions: pure geographic (RPG), purely environmental (RPE), mixed geographic-environmental (RPX), and unknown (RUN). These were calculated as follows: RPG = RT − RE; RPE = RT − RG; RMX = RE + RG − RT; and RUN = 1 − RT (cf. Borcard et al. 1992; Duivenvoorden et al. 2002; Tuomisto et al. 2003c). All multiple regressions on distance matrices were done using Permute 3.4 (Legendre et al. 1994).

### *Individual species level analyses*

To assess patterns in individual species, we used a series of regression models that include the effects of environmental and geographic distance variables on species abundances. Then, we used information-theoretic approach to determine which models provided the most explanatory power for the species abundance data (Burnham & Anderson 2002; Svenning et al. 2006). Two types of geographic distance variables were used (Bahn et al. 2006; Cressie 1993; Kaluzny et al. 1998; Lichstein et al. 2002; Svenning et al. 2006). First, the geographic trend variable measures how species abundance is related to geographic positions of the plots in the study area and thus influenced by broad-scale spatial trends. Second, the spatial autocorrelation variable measures how species abundance in one plot is related to abundances in neighboring plots.

We used ordinary least square regression (OLS) and conditional autoregressive models (CAR) models in a maximum likelihood framework to determine which combinations of environmental, geographic trend, and spatial autocorrelation variables best fit the patterns of abundance for individual species (Lichstein et al. 2002; Svenning et al. 2006). In species for which local spatial autocorrelation is important, a CAR model will show a large effects of spatial autocorrelation term with a corresponding reduction in either unexplained variance and/or the importance of the broad-scale trend or environmental variables. If spatial autocorrelation is not important, then OLS model will perform as well as the CAR models.

We restricted the analyses to 135 relatively common species, defined as having a mean density of at least one individual per hectare and being present in at least half of the 40 plots. The species abundance was square-root transformed to reduce skewness, resulting in only four of 135 species with skewness | > 1 |. For geographic trend variables (T), we used the UTM coordinates of the plots (X and Y) centered on their means and scaled to unit variance. We used the transformed values of the four most important environmental variables (E) from the community level analysis (Table 1) – the main fertility gradient, slope, pH and C:N ratio - which were centered and their variance scaled.

Conditional autoregressive (CAR) models are similar to OLS models, except they include a term in the explanatory variables to model the local spatial dependence in the residuals of the model. The CAR model is specified as $\mathbf{Y} = \mathbf{X}\beta + \rho\mathbf{C}(\mathbf{Y}-\mathbf{X}\beta) + \varepsilon$, where **Y** is the vector of observations, **X** is the matrix of explanatory variables, $\beta$ is the vector of regression coefficients, $\rho$ is the spatial autoregression parameter, **C** is a symmetric neighborhood matrix of weights ($w_{ij}$) that quantifies the degree to which the abundance of plot i is influenced by its neighbor j, and $\varepsilon$ is vector of random errors. The neighborhood matrix of $w_{ij}$ weights is based on the distance between neighbors and can take different forms depending on the nature of distance dependence in the system. We used the form $w_{ij} = 1/d_{ij}^k$ and tested models that allowed k to vary between 1 and 4. In the end, we used k=1 ($w_{ij} = 1/d_{ij}$) because it generated the highest overall the highest $r^2$ and $\rho$ values.

We initially developed a set of eight models, referred to below as the “trend/environment/rho” models, to explore the relative contribution of broad-scale geographic trend (X and Y together), all environment variables together and fine-scale spatial autocorrelation ($\rho$). The “trend/environment/rho” models included three OLS models that contained the following response variables but lacked the spatial autocorrelation term: intercept, geographic trend and environment ($M_{E+T}$); intercept and geographic trend ($M_T$); and intercept and environment ($M_E$). The three analogous CAR models - $M_{E+T+\ \rho}$, $M_{T+\ \rho}$ and $M_{E+\ \rho}$ - had the same combinations of explanatory variables plus the spatial autocorrelation parameter ρ. The final two models were the ρ-only CAR model (intercept + ρ) ($M_\rho$) and the null model with only an intercept term ($M_0$).

To explore the contribution of the four environmental variables individually and in all possible combinations, we also built an additional set of OLS-only models that included all permutations of geographic trend (X and Y together) and the four individual soil and topographic parameters. The outcome of this analysis, referred to below as the "all environmental variables" models, resulted in 32 models.

*Model assessment*

The relative support for each model within the two sets of models was determined using Akaike's Information Criterion (AIC), which takes into account both the relative likelihood (L) of each model and the model complexity, ie number of explanatory variables (k) included in the model (Akaike 1981). Because we had a low number of observations (n) but large number of explanatory variables (k), we used the small-sample bias-corrected form of the AIC, $AIC_c = -2\ln(L) + 2K + 2K(K + 1)/(n - K - 1)$. Because the OLS and CAR functions in S-plus used different error structures and methods to compute likelihoods, we used the approach of Svenning et al. (2006) to calculate the OLS likelihood estimates of the CAR models by adjusting the null CAR model (ρ=0) to equal the null OLS model. The model with the lowest $AIC_c$ among the OLS and adjusted CAR likelihoods was considered the best model ($AIC_c$min). The difference between the $AIC_c$ of each model and the best model ($\Delta AIC_c$) can be used to interpret the strength of support for alternative model compared to the best model. Models with $\Delta AIC_c \leq 2$ have strong support and should be considered along with the best model (Burnham & Anderson 2002).

Another approach to comparing models is Akaike weights, which can be interpreted as the probability that model i is the best model for the observed data. Akaike weights were calculated as: $W(i) = \exp(-1/2\Delta AIC_c(i)) / \Sigma \exp(-1/2\Delta AIC_c(i))$ where the denominator is summed over all models (Burnham & Anderson 2002). The $AIC_c$ weights were calculated for each species, then averaged over all species to produce a species-averaged Akaike weight. The probability that the best model includes a specific explanatory variable was determined by summing the species-averaged Akaike weights for all models that include that explanatory variable. For the OLS models of the "trend/environment/rho" model, we determined the amount of variance explained by each explanatory variable using the variance partitioning scheme described in the community analysis section of this paper.

When there is substantial support for a number of models, the relative importance of individual explanatory variables among the models can be assessed by computing weighted averages of the standardized regression coefficients. This procedure weights the regression coefficient for each parameter $\beta_i$ in each model by the AIC weight W(i) for that model, then sums over all models $\beta_{MA}=\Sigma W(i)\beta_i$. Standardized regression coefficients for each explanatory variable were determined for each species, then averaged over species to yield a species-average coefficient for each explanatory variable (Burnham & Anderson 2002).

For the “all environmental variables” models only, we tallied the number of common species that had significant relationships with each environmental variable using a Bonferroni-adjusted p-value. First, for each species, we determined if the individual environmental variables included in the best model were significant. Then for each environmental variable, we summed significant relationship over the species. The same procedure was done for models with strong support.

## Results

*Community composition*

Geographic distance was the most important variable shaping floristic similarity in the central Amazonian plots. A Mantel test between floristic similarity (Steinhaus’ index) and geographic distance yielded a relatively strong correlation (r= 0.57), which was mostly due to the sharp decline in floristic similarity between 0 - 5 km interplot distances (Fig. 1). Interplot floristic similarity was also correlated with interplot similarity in the soil and topographic variables (Table 1, Fig. 2). Half of the six soil and topographic variables (main fertility gradient, pH and C:N ratio) showed a significant correlation with floristic similarity using the simple Mantel tests. [Slope was not significant in the 40 plots, but was significant in the full 66 plots (Appendix 3).] Using partial Mantel tests to account for trends in floristic similarity related solely to geographic distance, only the main fertility gradient and C:N ratio were significant with correlation coefficients at 0.41 and 0.21, respectively (Table 1).

In the multiple regression analysis considering all the explanatory variables together, the final model contained interplot geographic distance and the main fertility gradient (Table 2, Appendix 4). Geographic distance had a higher standardized regression coefficient than the main fertility gradient. Partitioning the variance, pure geographic distance explained 26% of the community floristic variation, environmental factors (soil and topographic variables) explained 12%, and the interaction of distance and environment explained 7%, with 56% of the variation unexplained.

The plots in central Amazonia had lower overall beta diversity than did comparable plot networks in Peru, Ecuador, and Panama (Fig. 1). At small interplot distances (0-2.5 km), the three Amazonian networks had comparable floristic similarity between 0.45 and 0.52 fraction of shared species (the similarity of the Panama plots was much higher), but they diverged at larger distances (>2.5 km). At interplot distance 10 - 30 km, the similarity of the central Amazon plots remained fairly constant at 0.45, whereas the similarity within the Peruvian plots decreases steeply from 0.45 to 0.32. The floristic similarity of the Ecuador and Panama plots between 10 and 30 km was considerably lower than the central Amazon between 0.25 – 0.35 shared species.

*Individual species*

As in the community level analysis, geographic trend had somewhat greater importance than environmental variables in determining individual species abundances. Geographic trend was included in the best model for 41% of species compared to 30% for the environmental variables (Table 3, Appendix 5). However, considering all models with strong support, which are plausible alternatives to the best model, 89% of species contained the environmental variables compared to 59% that included the geographic trend variables (Table 3). The summed AIC weights, which is a measure of the weight of evidence that a variable should be included in the final model, were somewhat greater for broad-scale geographic trend (46%) than for the environmental variables (35%) (Table 3).

There was some, but not strong, support that spatial autocorrelation ($\rho$), ie fine-scale structuring, is important in this data set. The summed AIC weight for any model with $\rho$ was 37%, equal to that of models containing the environmental variables (Table 3). But spatial autocorrelation was included in only 15% of the best models (CAR models); 55% of best models had no rho parameter (OLS models) (Table 3). In 80-95% of species where a specific CAR models (for example, geographic trend + $\rho$) had strong support, the corresponding OLS model (for example, just geographic trend) also had strong support, suggesting that either a CAR or OLS model could be chosen for each species. The null model had quite strong support with an AIC weight of 20% (Table 3). For 30% of species, the null model was the best model, and for over half the species, the null model had strong support, meaning that a model that included at least one environmental or distance variable was only slightly better than a model with just an intercept.

The partitioning of the variance from the OLS models showed that 15% of the variance was explained by environmental variables, 8% by geographic trend and 8% by trend + environment. On average across species, the large majority of variance in species abundance (70%) was unexplained. For just the CAR models, the average variance explained ($r^2$) for the full model (RT) was 33%, thus the average unexplained variance (1-RT) was 67%, similar to the unexplained variance in the OLS models.

*Importance of individual environmental variables*

We used the "all environmental variables" OLS models to determine the relative importance of the four environmental variables for species abundances. Of the four environmental variables, the main fertility gradient and pH were more important in structuring species abundance than slope and C:N ratio. Both the main fertility gradient and pH were included in the best model for 39% of the species (Table 4, Appendix 5, 6). Also, the main fertility gradient and pH had higher $AIC_c$ weights than the other two environmental variables (Table 4). It is important to note, however, the data could not definitively arbitrate between the different models, as each environmental variable was included in at least one model that had strong support.

In the models with the best support, the main fertility gradient had a significant p-value for the most number of species (16%) compared to pH (9%), C:N ratio (3%) and slope (1%). Similarly, for models with strong support, the significant p-values were greatest for the main fertility gradient (21%), followed by pH (10%), slope (5%) and C:N ratio (4%). One quarter of the species had at least one environmental variable that was significant in the best model and 33% of the species had at least one significant environmental variable included in a model with strong support. There was evidence that certain species-rich families and genera have strong trends with particular environmental variables. The seven common Moraceae, 11 *Eschweilera* (Lecythidaceae), and nine *Protium* (Burseraceae) species showed a strong relationship with the main fertility gradient (c.f. Fine et al. 2004, who also found clay and sand specialists in the genus *Protium*), the six *Licania* (Chrysobalanceae) species with slope, and the five Myristicaceae species with C:N ratio (Appendix 5).

Finally, for model-averaged standardized regression parameters in the OLS models, the main fertility gradient and pH had the highest absolute coefficients averaged over all species (Table 5). Of the species with high coefficients (>0.10) for slope (32 species), 72% had positive coefficients, indicating that species tended to have greater abundance on steeper slopes. The main fertility gradient also had more species with positive than negative coefficients (Table 5). Of the species-rich taxa, *Licania* had high abundance at steeper slopes and *Eschweilera* tended to have higher abundance at sites with low clay content, and low amounts of nitrogen, carbon and cations (Appendix 5). pH and C:N ratio had more equal numbers of species with strong negative and positive relationships (Table 5).

We used the value of 34% as the number of species with a significant relationship with environmental variables to compare the results of this study to other studies. This was the percentage of species that had significant relationships with any environmental variables in the best or strongly-supported in the “all environmental variables” model. The 34% value also matches the percentage of species for which environment was included in the best model for the “trend/environment/rho” model (Table 3).

## Discussion

### *Community floristic similarity*

In our study, both geographic distance and environmental variables were important in determining community floristics patterns. The effect of geographic distance on floristic similarity was strongest at distances of 0-5 km, but then became relatively unimportant at distances of 5-30 km. A steep decline in similarity at local distances and a more gradual decline at larger distances has been reported throughout the Neotropics (Condit et al. 2002; Normand et al. 2006; Tuomisto et al. 2003c; Vormisto et al. 2004), and is consistent with the important role of dispersal limitation in determining beta diversity (Condit et al. 2002).

Of the soil and topographic variables, the main soil-fertility gradient influenced floristic similarity at all spatial scales, with widely separated sites on similar soils tending to be more similar floristically than are nearby sites on differing soils (see also Rankin-de Merona et al. 1992). The dominant soil trend in our central Amazonian landscape appears to be a gradient between clay-rich soils with higher organic matter, cation concentrations and nitrogen, and sand-dominated soils with opposite characteristics (Chauvel et al. 1987; Laurance et al. 1999; Richter & Babbar 1991). Comparable trends are found in other tropical forests where soil texture and cation concentrations strongly influence plant species composition (Phillips et al. 2003; Tuomisto et al. 2003b; Tuomisto et al. 2003a; Webb & Peart 2000). The C:N ratio of soils, which reflects soil nitrogen mineralization, also had a impact on community floristics. The relationship between C:N ratio and floristic composition has not been reported in other tropical forests, but can be important in temperate forests (Elgersman & Dhillion 2002; Schuster & Diekman 2005).

*Landscape and regional patterns*

The plots in central Amazonia exhibit two notable trends compared to the plot networks in northwestern and southwestern Amazonia: consistent levels of floristic similarity for plots >5 km apart and relatively low beta diversity (Condit et al. 2002;

Macia & Svenning 2005; Pitman et al. 2001). The differences in floristic similarity between the sites may result from different dispersal or speciation dynamics in central Amazon versus western Amazon forests (Condit et al. 2002). The uniformity in floristic similarity with increasing distance suggests that the central Amazonian landscape is dominated by a small group or “oligarchy” of common species (cf. Pitman et al. 2001). Similar to the western Amazon sites, 61% of the stems in central Amazonia are common species with densities of at least one stem per hectare. The oligarchy of common species on nutrient-starved soils of central Amazonia tends to be quite distinct from that of the geologically younger and richer soils of the western Amazon (Gentry 1990; Terborgh & Andresen 1998). Yasuní and Manu share 42 of their 150 commonest species (Pitman et al. 2001), but the central Amazon shares only 18 common species with Yasuni and 8 species with Manu. Although the common species of central Amazonia do not overlap with the common species in western Amazonia, we do not expect the common central Amazonian species to be specialist species localized to the BDFFP plots, as common species in other Neotropical forests tend to be widespread regionally and locally abundant (Duque et al. 2003; Eilu et al. 2004; Pitman et al. 2001).

### *Distance vs. environment in individual species*

Broad-scale geographic trend and the environmental variables were both important in determining individual species distributions. Spatial autocorrelation was less important, found in only 15% of the best models. The majority of the variance in individual species was not explained by geographic distance, environmental variables or spatial autocorrelation, indicating that importance of stochastic processes and possibly unmeasured environmental variables for species distributions.

Because of the high community floristic similarity at interplot distances less than 5 km, we expected distance variables to be important for individual species. Fine-scale spatial autocorrelation was less important and broad-scale trend more important than we expected. The lack of spatial autocorrelation may be due to the relatively large distances between plots. The minimum distance between plots was 100 m, whereas density dependent processes, such as dispersal limitations and Janzen-Connell effects, influence tree densities and dynamics at shorter spatial scales of 30 m or less (Condit et al. 2000, Hubbell et al. 2001, Queenborough et al. 2007).

The strong relationship between broad scale spatial trend and abundance for many species could have several underlying environmental or historical causes. The flora of the Manaus area is composed of species from several phytogeographic regions and many species are at the limit of their known range boundaries (Oliveira & Daly 1999). The expansion or contraction of these boundaries may leave geographic trends in some species' densities. The geographic trend in species abundance could result from unmeasured environmental variation. The western portion of the BDFFP is part of a different watershed and geological formation than the eastern portion. The road between Manaus and Caracas, which was constructed in the late 1960's, runs through the western portion of the study area, although not within 5 km of any plot. Although the wide disturbed right-of-way of this road could be a source of pioneer species, we believe that the likelihood is small that this factor could have affected the populations of trees ≥ 10 cm dbh in the study plots by the 1980s when our survey data were collected. Clues to the causes of the geographic variation might be obtained by investigating the species-rich families and genera that showed strong relationship geographic trend. In particular, the nine common *Protium* and 16 common Lecythidaceae species showed a strong relationships with geographic trend (Appendix 5).

### *Important environmental variables for species abundances*

As with community floristic similarity, the main fertility gradient was the most important environmental variable in relation to individual species abundances. However, pH was clearly the second most important environmental variable, in contrast to the community floristic similarity analysis in which C:N ratio was the second most important explanatory variable. Unlike cation concentration and soil texture, which seem to be widely important in influencing floristic patterns in tropical forests, the role of soil pH may be limited to certain regions or taxa. Soil pH has been found to have significant relationships with floristics in some tropical forests (Baillie 1987; Hall et al. 2004; Jones et al. 2006) but is unimportant in others (Paoli et al. 2006; Ruokolainen & Tuomisto 1998; Tuomisto et al. 2003a), including the Ducke Reserve in central Amazonia (Costa et al. 2005). Higher species abundances on soils of low pH might be expected due to the evolutionary origin of Amazonian flora of soils of low pH (Pärtel 2002). However, there were equal numbers of species with positive and negative species coefficients for pH, which does not lend support for greater species abundances at lower pH in this community.

It is important to note that for each species, each environmental variable was included in at least one model with strong support (Table 4), indicating that no single model definitively fit the data better than the other models. To narrow the candidates for the best fit model would require either more data or reducing the number of models under consideration (32 for the "all environmental variables" set of models). The analysis provided here can help guide the choice of the appropriate environmental variables for different species, genera and families for further study.

### *Meta-analyses across tropical forests*

The proportion of species having significant relationships with soil factors varied widely (25-82%) among studies of tropical forests (Fig. 3, Appendix 7). The methodologies of different studies varied tremendously, making it difficult to compare among them. For example, the number of plant taxa analyzed ranged from just a single genus (*Entandrophragma*) (Hall et al. 2004), a single family (Lauraceae, Dipterocarpaceae, Sterculiaceae) (Paoli et al. 2006; Sri-ngernyuang et al 2003; Yamada et al. 2006), and a single plant form (ferns or palms) (Svenning et al.1999; Tuomisto et al. 2002), to all trees (Cannon & Leighton 2004; Clark et al. 1998, 1999; Harms et al. 2001; Miyamoto et al. 2003; Phillips et al. 2003; Webb & Peart 2000). Despite these differences among studies (Appendix 7), one trend was apparent. The percentage of species with significant soil and topographic relationships was negatively correlated with the total number of species examined (Fig. 3). Among four studies that examined 100 or more species, all had a comparatively low percentage (25-55%) of species with significant soil or topographic relationships. The other ten studies had 55 or fewer species and the range of species with significant soil or topographic relationships was higher, between 65-82% for nine of the ten studies. Our study, in which 34% of the species were significantly related to at least one soil variable, was within the range of studies examining a large number of species. Other factors, such as the taxa, geographic region, soil variables or habitat type, study-area size, or total plot-area sampled, did not influence the proportion of identified species. It is possible that studies with a small number of species focus on species with previous observations of habitat specialization. Another possibility is that the studies with large numbers of species contain more rare species and fail to detect environmental relationships in the rare because of low sample size. The environmental variation within each study site was difficult to compare with the information provided in the literature, but would certainly be an important next step in a meta-analysis. To definitively study community floristics relationships with soil and topographic variables, more studies with large numbers of species and with soil data collected in a standardized manner are needed.

## Acknowledgements

We thank Hanna Tuomisto, Richard Condit, Ben Turner and four anonymous reviewers for commenting on the manuscript. Support was provided by the NASA-LBA Program, A. W. Mellon Foundation, Conservation, Food and Health Foundation, World Wildlife Fund-U.S., MacArthur Foundation, National Institute for Amazon Research, and Smithsonian Institution. This is publication number --- in the BDFFP technical series.

## References

Akaike, H. 1981. Modern development of statistical methods. in Eykhoff, P., (eds.) *Trends and progress in systems identification*. pp. 169-184, Pergamon Press, London.

Bahn, V., O'Connor, R. J. & Krohn, W. B. 2006. Importance of spatial autocorrelation in modeling bird distributions at a continental scale. *Ecography* 29:835-844.

Baillie, I. C., Ashton, P. S., Court, M. N., Anderson, J. A. R., Fitzpatrick, E. A. & Tinsley, J. 1987. Site characteristics and the distribution of tree species in mixed dipterocarp forest on tertiary sediments in Central Sarawak, Malaysia. *Journal of Tropical Ecology* 3:201-220.

Beinroth, F. H. 1975. Relationships between U.S. soil taxonomy, the Brazilian system, and FAO/UNESCO units. in Bornemisza, E. & Alvarado, A., (eds.) *Soil Management in Tropical America*. pp. 97-108, North Carolina State University, Raleigh, NC.

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

Brown, K. S. 1987. Soils and vegetation. in Whitmore, T. C. & Prance, G. T., (eds.) *Biogeography and Quaternary History in Tropical America*. pp. 19-45, Oxford Monographs in Biogeography, Oxford, UK.

Burnham, K. P. & Anderson, D. R. 2002. *Model selection and multi-model inference: a practical information-theoretic approach*. Springer-Verlag, New York.

Cannon, C. H. & Leighton, M. 2004. Tree species distributions across five habitats in a Bornean rain forest. *Journal of Vegetation Science* 15:257-266.

Chauvel, A., Lucas, Y. & Boulet, R. 1987. On the genesis of the soil mantle of the region of Manaus, Central Amazonia, Brazil. *Experientia* 43:234-240.

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

Clark, D. B., Clark, D. A. & Read, J. M. 1998. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. *Journal of Ecology* 86:101-112.

Clark, D. B., Palmer, M. W. & Clark, D. A. 1999. Edaphic factors and the landscape-scale distributions of tropical rain forest trees. *Ecology* 80:2662-2675.

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S. P., Foster, R. B., Itoh, A., LaFrankie, J. V., Lee, H. S., Losos, E., Manokaran, N., Sukumar, R. & Yamakura, T. 2000. Spatial patterns in the distribution of tropical tree species. *Science* 288:1414-1418.

Condit , R., Pitman, N., Leigh, E., Chave, J., Terborgh, J., Foster, R. B., Nunez, P., Aguilar, S., Valencia, R., Villa, G., Losos, E., Muller-Landau, H. & Hubbell, S. P. 2002. Beta diversity in tropical forest trees. *Science* 295:666-669.

Costa, F. R. C., Magnusson, W. E. & Luizao, R. C. 2005. Mesoscale distribution patterns of Amazonian understorey herbs in relation to topography, soil and watersheds. *Journal of Ecology* 93:863-878.

Cressie, N. A. C. 1993. *Statistics for spatial data*. John Wiley, New York.

Duivenvoorden, J. F., Svenning, J.-C. & Wright, S. J. 2002. Beta diversity in tropical forests. *Science* 295:636-637.

Duque, A., Cavelier, J. & Posada, A. 2003. Strategies of tree occupation at a local scale in terra firme forests in the Colombian Amazon. *Biotropica* 35:20-27.

Eilu, G., Hafashimana, D. L. N. & Kasenene, J. M. 2004. Tree species distribution in forests of the Albertine Rift, western Uganda. *African Journal of Ecology* 42:100-110.

Elgersman, A. M. & Dhillion, S. S. 2002. Geographical variability of relationships between forest communities and soil nutrients along a temperature-fertility gradient in Norway. *Forest Ecology and Management* 158:155-168.

Fearnside, P. M. & Leal-Filho, N. 2001. Soil and development in Amazonia: lessons from the Biological Dynamics of Forest Fragments Project. in Bierregaard, R. O., Gascon, C., Lovejoy, T. E. & Mesquita, R., (eds.) *Lessons from Amazonia: The Ecology and Conservation of a Fragmented Forest*. pp. 291-312, Yale University Press, New Haven, CT.

Fine, P. V. A., Mesones, I. & Coley, P. D. 2004. Herbivores promote habitat specialization by trees in Amazonian forests. *Science* 305:663-665.

Gentry, A. H. 1990. Floristic similarities and differences between Southern Central America and Upper and Central Amazonia. in Gentry, A. H., (eds.) *Four Neotropical Rainforests*. pp. 141-157, Yale University Press, New Haven, CT.

Hall, J. S., McKenna, J. J., Ashton, P. M. S. & Gregoire, T. G. 2004. Habitat characterizations underestimate the role of edaphic factors controlling the distribution of Entandrophragma. *Ecology* 85:2171-2183.

Harms, K. E., Condit, R., Hubbell, S. P. & Foster, R. B. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *Journal of Ecology* 89:947-959.

Hubbell, S. P. 2001. *The Unified Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, New Jersey, USA.

Hubbell, S. P., Ahumada, J. A., Condit, R. & Foster, R. B. 2001. Local neighborhood effects on long-term survival of individual trees in a neotropical forest. *Ecological Research* 16:859-875.

Jones, M. M., Tuomisto, H., Clark, D. B. & Olivas, P. 2006. Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation in rain forest ferns. *Journal of Ecology* 94:181-195.

Kaluzny, S. P., Vega, S. C., Cardoso, T. P. & Shelly, A. A. 1998. *S+ Spatial Stats User's Manual*. MathSoft, Seattlle.

Laurance, W. F., Fearnside, P. M., Laurance, S. G., Delamonica, P., Lovejoy, T. E. & Rankin, d. 1999. Relationship between soils and Amazon forest biomass: a landscape-scale. *Forest Ecology and Management* 118:1-3.

Laurance, W. F., Lovejoy, T. E., Vasconcelos, H. L., Bruna, E., Didham, R., Stouffer, P., Gascon, C., Bierregaard, R., Laurance, S. & Sampiao, E. 2002. Ecosystem decay of Amazonian forest fragments: a 22-year investigation. *Conservation Biology* 16:605-618.

Laurance, W. F., Oliveira, A. A., Laurance, S. G., Condit, R., Dick, C. W., Andrade, A., Nascimento, H. E. M. & Lovejoy, T. E. 2005. Altered tree communities in undisturbed Amazonian forests: A consequence of global change? *Biotropica* 37:160-162.

Legendre, P., Lapointe, F.-J. & Casgrain, P. 1994. Modeling brain evolution from behaviour: a permutational regression approach. *Evolution* 48:1487-1499.

Legendre, P. & Legender, L. 1998. *Numerical Ecology*. Elsevier, Amsterdam.

Lichstein, J. W., Simons, T. R., Shriner, S. A. & Franzreb, K. E. 2002. Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 72:445-463.

Lovejoy, T. E., Bierregaard, R. O., Rylands, A. B., Malcolm, J. R., Quintela, C. E., Harper, L. H., Brown, K. S., Powell, A. H., Powell, G. V. N., Schubart, H. O. R. & Hays, M. B. 1986. Edge and other effects of isolation on Amazon forest fragments. in Soule, M. E., (eds.) *Conservation Biology: The Science of Scarcity and Diversity*. pp. 257-285, Sinauer, Sunderland, MA.

Macia, M. J. & Svenning, J. C. 2005. Oligarchic dominance in western Amazonian plant communities. *Journal of Tropical Ecology* 21:613-626.

Miyamoto, K., Suzuki, E., Kohyama, T., Seino, T., Mirmanto, E. & Simbolon, H. 2003. Habitat differentiation among tree species with small-scale variation of humus depth and topography in a tropical heath forest of Central Kalimantan, Indonesia. *Journal of Tropical Ecology* 19:43-54.

Nelson, B. W. 2005. Pervasive alteration of tree communities in undisturbed Amazonian forests. *Biotropica* 37:158-159.

Nelson, B. W., Kapos, V., Adams, J. B., Oliveira, W. J., Braun, O. P. g. & Amaral, I. L. 1994. Forest disturbance by large blowdowns in the Brazilian Amazon. *Ecology* 75:853-858.

Normand, S., Vormisto, J., Svenning, J.-C., Grandez, C. & Balslev, H. 2006. Geographical and environmental controls of palm beta diversity in paleo-riverine terrace forests in Amazonian Peru. *Plant Ecology* 186:161-176.

Oliveira, A. A. & Daly, D. C. 1999. Geographic distribution of tree species occurring in the region of Manaus, Brazil: implications for regional diversity and conservation. *Biodiversity and Conservation* 8:1245-1259.

Oliveira, A. A. & Mori, S. A. 1999. A central Amazonian terra firme forest. I. High tree species richness on poor soils. *Biodiversity and Conservation* 8:1219-1244.

Palmer, M. W. 2006. Distance decay in an old-growth neotropical forest. *Journal of Vegetation Science* 16:161-166.

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

Pärtel, M. 2002. Local plant diversity patterns and evolutionary history at the regional scale. *Ecology* 83:2361-2366.

Phillips, O. L., Nunez Vargas, P., Monteagudo, A. L., Pena Cruz, A., Chuspse Zans, M.-E., Galiano Sanchez, W., Li-Halla, M. & Rose, S. 2003. Habitat association among Amazonian tree species: a landscape-approach. *Journal of Ecology* 91:757-775.

Piperno, D. R. & Becker, P. 1996. Vegetational history of a site in the Central Amazon basin derived from phytolith and charcoal records from natural soils. *Quarternary Research* 45:202-209.

Pitman, N. C. A., Terborgh, J. W., Silman, M. R., Nunez Vargas, P., Neill, D. A., Ceron, C. E., Palacios, W. A. & Aulestia, M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* 82:2102-2117.

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.

Queenborogh, S. A., Burslem, D. F. R. P., Garwood, N. C. & Valencia, R. 2007. Neighborhood and community interactions determine the spatial pattern of tropical seedling survivial. *Ecology* 88:2248-2258.

Rankin-de Merona, J. M., Prance, J. M., Hutchings, R. W., Silva, M. F., Rodrigues, W. A. & Uehling, M. A. 1992. Preliminary results of a large scale inventory of upland rain forest in the central Amazon. *Acta Amazonica* 22:493-534.

Richter, D. D. & Babbar, K. I. 1991. Soil diversity in the tropics. *Advances in Ecological Research* 21:315-389.

Ruokolainen, K. & Tuomisto, H. 1998. Vegetacion natural de la zona de Iquitos. in Kalliola, R. & Flores Paitan, S., (eds.) *Geoecologia y desarrollo amazonico: estudio integrado en la zona de Iquitos, Peru*. pp. 253-365.

Santos, G. M., Gomes, P. R. S., Anjos, R. M., Cordeiro, R. C., Turcq, B. J., Sifeddine, A., de Tada, M. L., Creswell, R. G. & Fifield, L. K. 2000. $^{14}$C AMS dating of fires in the central Amazon rain forest. *Nuclear Instruments and Methods in Physics Research B* 172:761-766.

Schuster, B. & Diekmann, M. 2005. Species richness and environmental correlates in deciduous forests of Northwest Germany. *Forest Ecology and Management* 206:197-205.

Smouse, P. E., Long, J. C. & Sokal, R. R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* 35:627-632.

Sri-Ngernyuang, K., Kanzaki, M., Mizuno, T., Noguchi, H., Teejuntuk, S., Sungpalee, C., Hara, M., Yamakura, T., Sahunalu, P., Dhanmanonda, P. & Bunyavejchewin, S.

2003. Habitat differentiation of Lauraceae species in a tropical lower montane forest in northern Thailand. *Ecological Research* 18:1-14.

Svenning, J. C. 1999. Microhabitat specialization in a species-rich palm comunit in Amazonian Ecuador. *Journal of Ecology* 87:55-65.

Svenning, J.-C., Engelbrecht, B. M. J., Kinner, D. A., Kursar, T. A., Stallard, R. F. & Wright, S. J. 2006. The relative roles of environment, history and local dispersal in controlling the distributions of common tree and shrub species in a tropical forest landscape, Panama. *Journal of Tropical Ecology* 22:575-586.

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

Terborgh, J. & Andresen, E. 1998. The composition of Amazonian forests: patterns at local and regional scales. *Journal of Tropical Ecology* 14:645-664.

Tuomisto, H., Poulsen, A. D., Ruokolainen, K., Moran, R. C., Quintana, C., Celi, J. & Canas, G. 2003a. Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. *Ecological Applications* 13:352-371.

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

Tuomisto, H., Ruokolainen, K., Poulsen, A. D., Moran, R. C., Quintana, C., Canas, G. & Celi, J. 2002. Distribution and diversity of pteridophytes and Melastomataceae along edaphic gradients in Yasuni National Park, Ecuadorian Amazonia. *Biotropica* 34:516-533.

Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. 2003c. Dispersal, environment, and floristic variation of Western Amazonian forests. *Science* 299:241-244.

Valencia, R., Foster, R. B., Villa, G., Condit, R., Svenning, J. C., Hernandez, C., Romoleroux, K., Losos, E., Magard, E. & Balslev, H. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *Journal of Ecology* 92:214-229.

Vormisto, J., Svenning, J. C., Hall, P. & Balslev, H. 2004. Diversity and dominance in palm (Arecaceae) communities in terra firme forests in the western Amazon basin. *Journal of Ecology* 92:577-588.

Webb, C. O. & Peart, D. R. 2000. Habitat associations of trees and seedlings in a Bornean rain forest. *Journal of Ecology* 88.

Yamada, T., Tomita, A., Itoh, A., Yamakura, T., Ohkubo, T., Kanzaki, M., Tan, S. & Ashton, P. S. 2006. Habitat associations of Sterculiaceae trees in a Bornean rain forest plot. *Journal of Vegetation Science* 17:559-566.

Table 1. Results of simple and partial Mantel tests between floristic similarity (as measured by Steinhaus' index), interplot differences in environmental variables and interplot geographic distance in central Amazonia. The environmental variables and distance have been transformed. The Mantel correlation coefficient (*r*) and significance level (*p*) is shown for each test. Relationships in bold are statistically significant at a Bonferroni-adjusted level of p<0.007.

| Independent Variable | Transformation | Variable ~ Distance | | Floristic Similarity ~ Variable | | Floristic Similarity ~ Variable \| Distance | |
|---|---|---|---|---|---|---|---|
| | | *r* | *p* | *r* | *p* | *r* | *p* |
| Geographic distance | log10(x) | -- | -- | **0.57** | **0.0005** | -- | -- |
| Main fertility gradient | $x^2$ | **0.17** | **0.0020** | **0.43** | **0.0005** | **0.41** | **0.0005** |
| pH | x | **0.41** | **0.0005** | **0.31** | **0.0005** | 0.10 | 0.0640 |
| Slope | $x^{0.5}$ | 0.06 | 0.0645 | 0.15 | 0.0220 | 0.14 | 0.0340 |
| C:N ratio | $-x^{-3}$ | **0.19** | **0.0005** | **0.28** | **0.0005** | **0.21** | **0.0030** |
| Water | $x^{0.5}$ | 0.06 | 0.0630 | 0.17 | 0.0090 | 0.16 | 0.0115 |
| Phosphate | $x^2$ | 0.09 | 0.0240 | 0.11 | 0.0445 | 0.07 | 0.1280 |

Table 2. Standardized regression coefficients from multiple regression analyses with interplot floristic similiary (Steinhaus index) as the response variable and interplot geographic distance and interplot variation in six environmental variables as the explanatory variables. A Bonferroni-corrected p-value of 0.005 was used in forward and backward selection to derive the final model from the full model.

| | Full model | Final model |
|---|---|---|
| | Coefficient (p) | Coefficient (p) |
| Full model | | |
| Distance | 0.43 (<0.0001) | 0.52 (0.0010) |
| Main fertility gradient | 0.33 (<0.0001) | 0.34 (0.0010) |
| pH | 0.14 (0.0040) | |
| C:N ratio | 0.09 (0.0340) | |
| Water | 0.09 (0.0420) | |
| Slope | 0.05 (0.1690) | |
| Phosphate | 0.10 (0.0190) | |
| $r^2$ | 0.49 (0.0001) | 0.44 (0.0005) |

Table 3. Comparison of “trend/environment/rho” models with all possible combinations of broad-scale geographic trend, environmental variables (main fertility gradient, slope, pH and C:N ratio together) and fine-scale spatial autocorrelation (ρ). The $AIC_c$ weights were averaged over all species. The $AIC_c$ weights can be interpreted as the probability that the model is the best model in this set for the observed data.

| Model | All models $AIC_c$ weights | Best Model # spp | Best model or strong support # spp | Best model $r^2$ |
|---|---|---|---|---|
| Geographic trend + environment | 8% | 9 (7%) | 37 (27%) | 0.46 |
| Geographic trend | 22% | 41 (30%) | 97 (72%) | 0.23 |
| Environment | 14% | 24 (18%) | 65 (48%) | 0.34 |
| Geographic trend + environment + ρ | 5% | 1 (1%) | 30 (22%) | NA |
| Geographic trend + ρ | 12% | 7 (5%) | 48 (36%) | 0.46 |
| Environment + ρ | 8% | 5 (4%) | 76 (56%) | 0.45 |
| P | 12% | 7 (5%) | 71 (53%) | 0.18 |
| Null | 20% | 41 (30%) | 75 (56%) | NA |
| Geographic trend in any model | 46% | 56 (41%) | 79 (59%) | 0.30 |
| Environment in any model | 35% | 41 (30%) | 120 (89%) | 0.39 |
| ρ in any model (CAR) | 37% | 20 (15%) | 135 (100%) | 0.37 |
| No ρ in model (OLS) | 44% | 74 (55%) | 133 (99%) | 0.30 |

Table 4. Summary of results from "all environmental variables" ordinary least squares (OLS) models, which includes all possible combinations of geographic trend and the four environmental variables as explanatory variables resulting in 32 models. For each species, AIC$_c$ weights were calculated as the sum of AIC$_c$ weights from all models that include the variable. Then the AIC$_c$ weight for each variable was averaged over the species. Individual model results are given in Appendix A2.

| Model | Sum of AIC$_c$ weights | Best model | Best model or strong support |
|---|---|---|---|
| | | # spp | # spp |
| Null | 5.1% | 18 (13%) | 58 (43%) |
| Main fertility gradient in any model | 49.0% | 52 (39%) | 135 (100%) |
| Slope in any model | 39.8% | 33 (24%) | 135 (100%) |
| pH in any model | 46.2% | 53 (39%) | 135 (100%) |
| C:N ratio in any model | 36.2% | 20 (15%) | 134 (99%) |
| Geographic trend in any model | 43.5% | 49 (36%) | 134 (99%) |
| Environmental variable in any model | 89.6% | 104 (77%) | 135 (100%) |

Table 5. Weighted standardized regression coefficients for individual environmental variables averaged across all species. To get these values, the regression coefficient for each standardized parameter $\beta_i$ in each model is weighted by the AIC weight W(i) for that model, then summed over all models $\beta_{MA}=\Sigma W(i)\beta_i$. The models used to calculate these coefficients was the “all environmental variables” models, which included all possible combinations of one geographic trend and four environmental variables for a total of 32 models (Appendix 5). The coefficients can be positive or negative, depending on the slope of the relationship between species abundance and the environmental variable. The mean is derived from the absolute value of the coefficient.

| | | No. of species with coefficients > 0.10 | |
|---|---|---|---|
| Environmental Variable | Mean | Positive slope | Negative slope |
| Main fertility gradient | 0.141 | 34 (25%) | 20 (15%) |
| Slope | 0.069 | 23 (17%) | 9 (7%) |
| pH | 0.112 | 32 (24%) | 24 (18%) |
| C:N ratio | 0.059 | 12 (9%) | 11 (8%) |

## FIGURE CAPTIONS

Figure 1. Floristic similarity (measured with Sørenson's index) as a function of geographic distance between plots in central Amazonia and three other Neotropical plot networks. Data for Panama, Peru, and Ecuador were adapted from Condit *et al*. (2002). The data points are for central Amazonia only. The lines were generated by classifying the floristic similarity data based on clusters of interplot distances and taking the average in each distance class.

Figure 2. Floristic similarity (measured with Steinhaus' index) as a function of difference in percent clay (an indicator of the main fertility gradient) between 40 plots in central Amazonia. Inset is a plot of floristic similarity versus different in C:N ratio between plots. The lines are based on a linear regression.

Figure 3. Comparison of the percentage of individual species with significant relationships with soil and topographic variables versus the total number of species for 14 studies in tropical forests. Squares indicate studies where individual soil or topographic variables were used. Circles indicate studies where plots were divided into two or more distinct habitats. "X" indicates that the species in the study were only drawn from one or two particular taxa, such as only Dipterocarpaceae, or only palms. Information on the studies can be found in Appendix 7.

Fig. 1

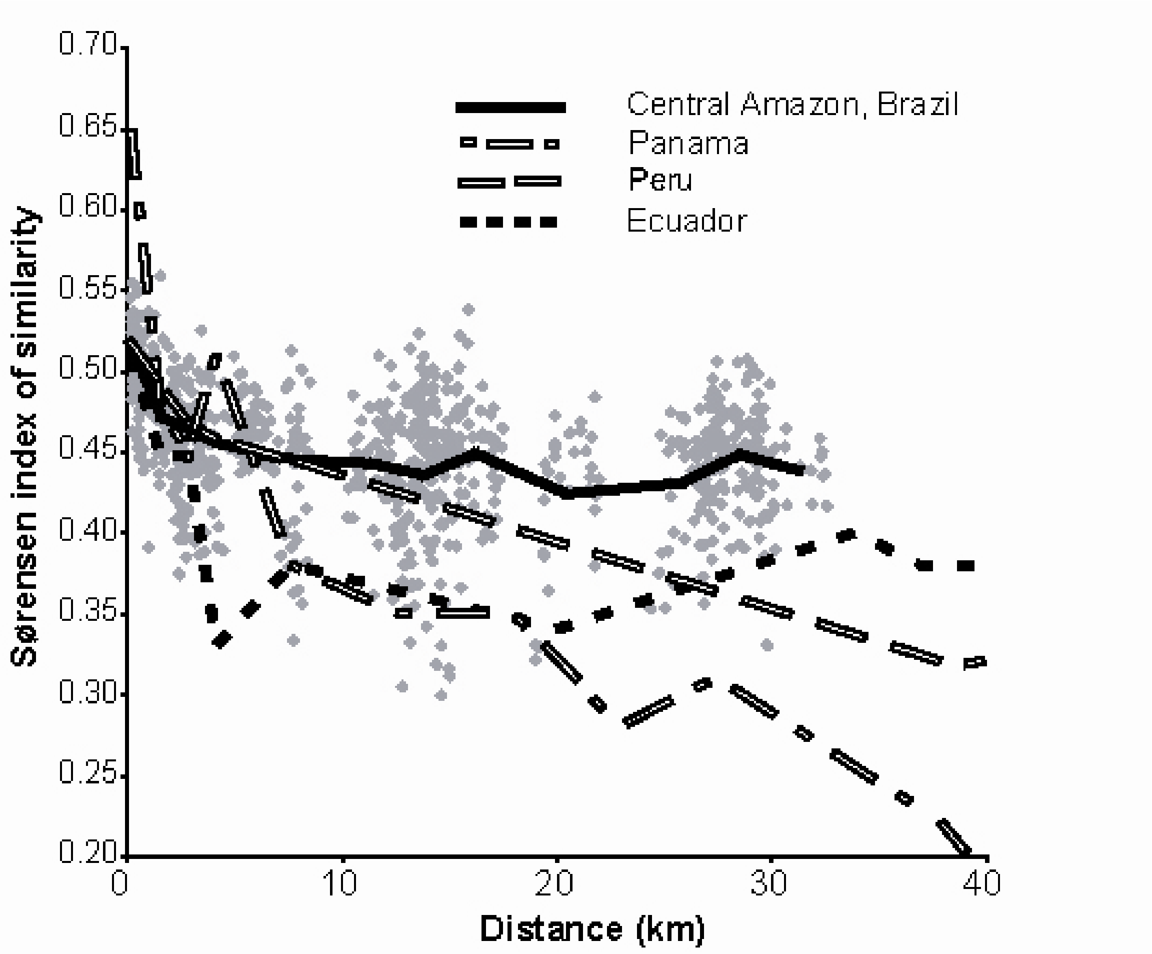

Central Amazon, Brazil
Panama
Peru
Ecuador
Sørensen index of similarity
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0
10
20
30
40
Distance (km)

Fig. 2

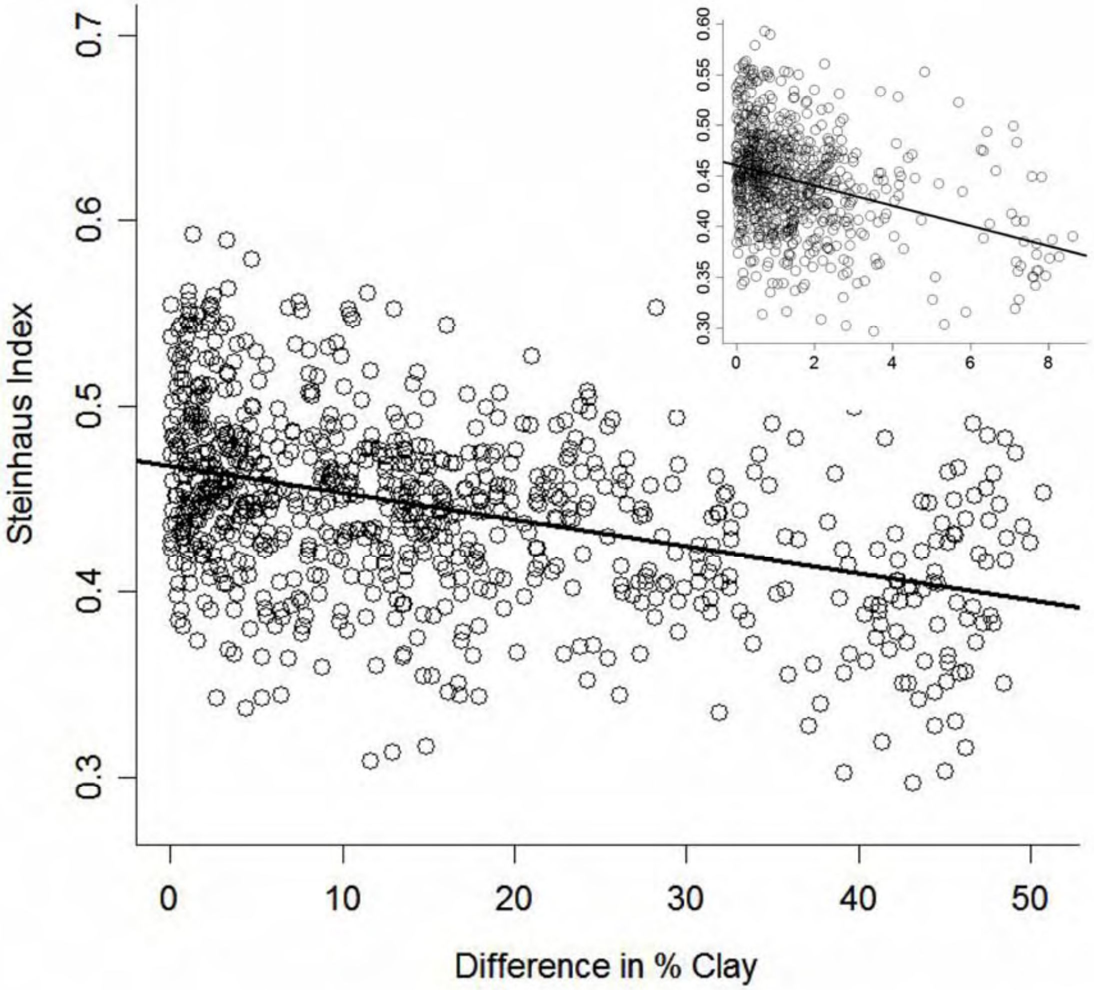
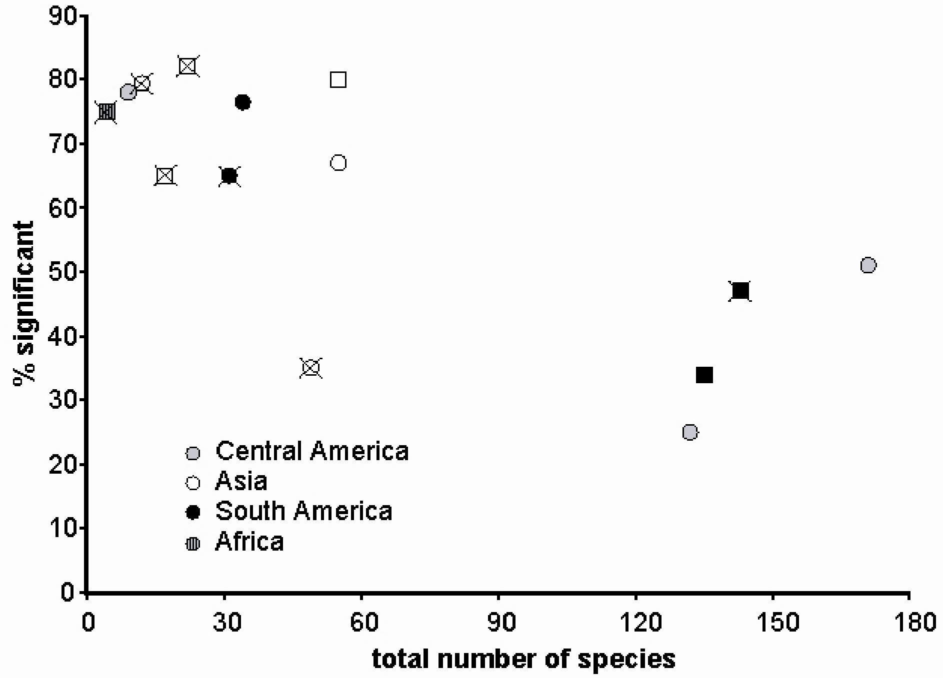

Steinhaus Index
0.3
0.4
0.5
0.6
0.7
Difference in % Clay
0
10
20
30
40
50
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0
2
4
6
8

Fig. 3



% significant
90
80
70
60
50
40
30
20
10
0
0
30
60
90
120
150
180
total number of species
● Central America
○ Asia
● South America
◍ Africa

Appendix 1. Laboratory methods for soil analysis.

Composite samples for each quadrat were oven-dried for 24 hours at 105°C, cleaned by removing stones and charcoal fragments, and then passed through 20 mm and 2 mm sieves. In all cases, values for soil parameters were derived separately for each quadrat, then combined to yield a mean value for each 1-ha plot.

Textural analyses were conducted to determine percentages of clay (particles <0.002 mm diameter), silt (0.002-0.05 mm), fine sand (0.05-0.2 mm), and coarse sand (0.2-2 mm), using the pipette method. Available water capacity (equivalent to "maximum plant-available water"), a measure of the amount of water the soil can hold in a form extractable by plant roots, was estimated as the difference between the field capacity (moisture content retained in soil under a suction of 0.33 atmospheres) and the wilting point (moisture content retained at 15 atmospheres), using a pressure membrane apparatus. As is standard practice, samples were dried, sieved, and re-wetted before determining available water capacity, making the results only an index of water available to plants in the field.

A pH meter was used to measure soil pH in water with a ratio of 1 ml soil to 25 ml distilled water (SCNLS-EMBRAPA 1979). Total N was determined by Kjeldahl digestion (Parkinson & Allen 1975). Total C was determined by dry combustion and organic C by the Walkley-Black method. Total P was determined by digestion in $HNO_3^+$, $HClO_4$, and HF (Lim & Jackson 1982), and reaction with ammonium molybdate. Extractable nutrients were determined as follows, using separate sub-samples for each group of analyses: $PO_4^{3-}$ was measured in an autoanalyzer using the molybdenum blue method (Jorgenson 1977); $K^+$ was determined by atomic emission spectroscopy at the Brazilian Center for Nuclear Energy in Agriculture (CENA), Piracicaba, São Paulo, while $Ca^{2+}$, $Mg^{2+}$, $Na^+$, S, Cu, Fe, $Mn^{2+}$, $Zn^+$, $Al^{3+}$, and $H^+$ in soil were measured by atomic absorption spectrophotometry at CENA after digestion in $HClO_4$, $HNO_3$, and $H_2SO_4$ (Jorgenson 1977).

Cation measures were derived without $Na^+$, which is generally a minor constituent of exchangeable bases and was recorded at only a limited number of study sites. These measures included cation exchange capacity (CEC), which is the sum of $K^+$, $Ca^{2+}$, $Mg^{2+}$, $Al^{3+}$, and $H^+$; total exchangeable bases (TEB), the sum of $K^+$, $Ca^{2+}$, $Mg^{2+}$, and $Na^+$; and aluminum saturation, which is $((Al^{3+} + H^+)/CEC) X 100$. Samples were digested in $HClO_4$, $HNO_3$, and $H_2SO_4$, and extracts for cation determination were buffered to pH 7.0, the standard practice in Brazil (SNLCS-EMBRAPA 1979). A ratio of 1 ml soil to 10 ml solution was used, and the mixture was allowed to decant overnight without filtration (SNLCS-EMBRAPA 1979).

Organic carbon to total nitrogen (C:N) ratios were calculated to provide an index of N mineralization. Low C:N values indicate recalcitrant N, values in the 10-12 range are considered ideal for plant growth, while high values indicate inhibited nitrification, often due to waterlogging (Young 1976).

## References

Jorgenson, S. S. 1977. *Guia Analitico: Metodologia Utilizada para Analises Quimicas de Rotina*. Centro de Energia Nuclear na Agricultura (CENA), Piracicaba, Sao Paulo, Brazil.

Lim, C. H. & Jackson, M. L. 1982. Dissolution for total elemental analysis. In: Page, A. L., Miller, R. H. & Keeney, D. R., (eds.) *Methods of Soil Analysis, Part 2: Chemical and Microbiological Properties*. pp. 1-12. American Society of Agronomy and Soil Science Society of America., Madison, WI.

Parkinson, J. A. & Allen, S. E. 1975. A wet oxidation procedure suitable for the determination of nitrogen and mineral nutrients in biological material. *Communications in Soil Science and Plant Analysis* 6:1-11.

SCNLS-EMBRAPA. 1979. *Manual de Metodos de Analise de Solo. Servicio Nacional de Levantamento e Conservacao de Solos-Empresa Brasileira de Pesquisa Agropecuaria*. SNCLS-EMBRAPA, Rio de Janeiro, Brazil.

Young, A. 1976. *Soils and Soil Survey*. Cambridge University Press, Cambridge.

Appendix 2. Pearson correlations between the first two PCA axes and different soil and topographic variables in central Amazonia. Soil and topographic variables are listed in decreasing strength of association with PC1, and variables in boldface were used to assess the distributions of individual tree species.

| Soil Variable | Units | Transfor-mation | Min | Max | Mean | Pearson correlations | |
|---|---|---|---|---|---|---|---|
| | | | | | | PC1 | PC2 |
| **Clay Fraction** | **%** | **$x^{2}$** | **18.0** | **73.3** | **53.6** | **0.97** | **-0.04** |
| Fine Sand Fraction | % | log x | 1.3 | 18.1 | 5.7 | -0.95 | 0.02 |
| Total Sand Fraction | % | log x | 5.8 | 71.4 | 25.7 | -0.95 | 0.09 |
| Coarse Sand Fraction | % | log x | 4.4 | 56.6 | 20.0 | -0.94 | 0.12 |
| Total Nitrogen | % | $x^{2}$ | 0.08 | 0.22 | 0.16 | 0.94 | -0.15 |
| Total Exchangeable Bases | m.e./100g dry soil | log x | 0.064 | 0.303 | 0.192 | 0.91 | 0.24 |
| Magnesium | m.e./100g dry soil | x | 0.013 | 0.125 | 0.073 | 0.90 | 0.10 |
| Zinc | ppm | $x^{2}$ | 0.63 | 2.99 | 1.65 | -0.89 | -0.19 |
| Aluminum Saturation | % of CEC | $-x^{-1}$ | 87.8 | 96.2 | 92.5 | -0.79 | 0.00 |
| Total Organic C | % | $x^{2}$ | 1.0 | 2.0 | 1.6 | 0.78 | 0.15 |
| Sodium | m.e./100g dry soil | $x^{0.5}$ | 0.017 | 0.106 | 0.051 | 0.77 | -0.14 |
| Manganese | ppm | $x^{2}$ | 0.87 | 2.49 | 1.73 | 0.76 | -0.11 |
| H | m.e./100g dry soil | $x^{2}$ | 0.393 | 0.841 | 0.681 | 0.73 | -0.16 |
| Calcium | m.e./100g dry soil | $x^{0.5}$ | 0.015 | 0.131 | 0.063 | 0.71 | 0.57 |
| Cation Exchange Capacity (CEC) | m.e./100g dry soil | $x^{3}$ | 1.07 | 3.31 | 2.44 | 0.69 | 0.57 |
| Silt Fraction | % | x | 8.3 | 32.8 | 20.6 | 0.66 | -0.21 |
| Iron | ppm | $x^{2}$ | 77 | 185 | 128 | 0.66 | 0.00 |
| Potassium | | $x^{0.5}$ | 0.028 | 0.077 | 0.057 | 0.65 | -0.51 |
| **Slope** | **%** | $x^{0.5}$ | **0.3** | **37.6** | **12.3** | **-0.51** | **-0.08** |
| Aluminum | m.e./100g dry soil | $x^{0.5}$ | 0.42 | 2.22 | 1.57 | 0.51 | 0.75 |
| Total Phosphorus | | x | 50.9 | 177.2 | 114.0 | 0.48 | -0.81 |
| Sulfur | ppm | $x^{2}$ | 11.3 | 14.9 | 12.9 | 0.46 | 0.33 |
| Copper | ppm | x | 0.10 | 0.54 | 0.30 | 0.41 | 0.05 |
| **Available Water Capacity** | **% by weight** | **$x^{0.5}$** | **3.3** | **11.3** | **7.2** | **0.39** | **-0.38** |
| **C:N Ratio** | **dimensionless** | **$-x^{-3}$** | **8.3** | **16.6** | **10.0** | **-0.20** | **0.47** |
| **Phosphate** | **m.e./100g dry soil** | **$x^{2}$** | **0.014** | **0.041** | **0.028** | **-0.05** | **-0.24** |
| **pH** | **pH units** | **x** | **3.7** | **4.5** | **4.1** | **0.05** | **-0.92** |

Appendix 3. Results of simple and partial Mantel tests between floristic similarity, interplot differences in environmental variables and interplot geographic distance in central Amazonia. The environmental variables and distance have been transformed. Floristic similarity was measured as either 1-Sørenson's index or Steinhaus' index. If the distance or environmental variable was available for more than 40 plots, the results for all plots containing that variable are listed in addition to the results for 40 plots. The results for Steinhaus' index and 40 plots are in the main text. The Mantel correlation coefficient (*r*) and significance level (*p*) is shown for each test. Relationships in bold are statistically significant at a Bonferroni-adjusted level of $p<0.007$.

| Floristic Similarity Index | Independent Variable | Transformation | No. of plots | Variable ~ Distance |  | Floristic Similarity ~ Variable |  | Floristic Similarity ~ Variable \| Distance |  |
|---|---|---|---|---|---|---|---|---|---|
|  |  |  |  | *r* | *p* | *r* | *p* | *r* | *p* |
| Sørenson’s | Geographic distance | log10(x) | 66 | -- | -- | **0.53** | **0.0010** | -- | -- |
|  |  |  | 40 | -- | -- | **0.57** | **0.0005** | -- | -- |
|  | Main fertility gradient | $x^2$ | 66 | **0.12** | **0.0035** | **0.30** | **0.0005** | **0.28** | **0.0005** |
|  |  |  | 40 | **0.17** | **0.0020** | **0.42** | **0.0005** | **0.40** | **0.0005** |
|  | pH | $x$ | 40 | **0.41** | **0.0005** | **0.33** | **0.0005** | 0.13 | 0.0305 |
|  | Slope | $x^{0.5}$ | 66 | **0.11** | **0.0025** | **0.24** | **0.0010** | **0.22** | **0.0005** |
|  |  |  | 40 | 0.06 | 0.0645 | 0.15 | 0.0150 | 0.14 | 0.0305 |
|  | C:N ratio | $-x^{-3}$ | 40 | **0.19** | **0.0005** | **0.31** | **0.0005** | **0.25** | **0.0005** |
|  | Water | $x^{0.5}$ | 40 | 0.06 | 0.0610 | 0.11 | 0.0500 | 0.09 | 0.0715 |
|  | Phosphate | $x^2$ | 40 | 0.09 | 0.0265 | 0.04 | 0.2810 | 0.00 | 0.5790 |
|  | PC axis 1 | $x$ | 40 | **0.18** | **0.0020** | **0.42** | **0.0005** | **0.39** | **0.0005** |
| Steinhaus’ | Geographic distance | log10(x) | 65 | -- | -- | **0.55** | **0.0005** | -- | -- |
|  | Main fertility gradient | $x^2$ | 65 | **0.12** | **0.0020** | **0.29** | **0.0005** | **0.26** | **0.0005** |
|  | Slope | $x^{0.5}$ | 66 | **0.11** | **0.0030** | **0.23** | **0.0010** | **0.20** | **0.0005** |
|  | PC axis 1 | $x$ | 40 | **0.18** | **0.0010** | **0.42** | **0.0005** | **0.39** | **0.0005** |

Appendix 4. Standardized regression coefficients from multiple regression analyses with interplot floristic similarity as the response variable and interplot geographic distance and interplot variation in six environmental variables as the explanatory variables. Floristic similarity is calculated as both 1-Sørensons' index and Steinhaus' index. Results for Steinhaus' index for 40 plots is shown in table 2 in the text. Only values for the main fertility gradient and slope were available for all 66 plots. A Bonferroni-corrected p-value of 0.005 was used in forward and backward selection to derive the final model from the full model.

| | Sørenson's<br>40 plots<br>Full model<br>Coefficient (p) | Sørenson's<br>40 plots<br>Final model<br>Coefficient (p) | Sørenson's<br>66 plots<br>Final model<br>Coefficient (p) | Steinhaus'<br>66 plots<br>Final model<br>Coefficient (p) |
|---|---|---|---|---|
| Full model | | | | |
| Distance | 0.42 (<0.0001) | 0.49 (0.0010) | 0.49 (<0.001) | 0.53 (<0.001) |
| Main fertility gradient | 0.31 (<0.0001) | 0.30 (0.0010) | 0.21 (<0.001) | 0.22 (<0.001) |
| pH | 0.15 (0.0030) | | | |
| C:N ratio | 0.15 (0.0030) | 0.15 (0.0020) | | |
| Water | 0.02 (0.3450) | | | |
| Slope | 0.06 (0.1280) | | 0.14 (0.001) | |
| Phosphate | 0.03 (0.0254) | | | |
| $r^2$ | 0.48 (0.0001) | 0.45 (0.0005) | 0.36 (0.0005) | 0.35 (0.0005) |

Appendix 5. Results of ordinary least squares (OLS) and conditional autoregressive (CAR) models for individual species abundances. The "trend/env/rho" consists of 8 models that are all possible combinations of geographic trend, environmental variables (main fertility gradient, slope, pH and C:N ratio together) and a spatial autocorrelation parameter (ρ). The "all env variables" consists of 32 models that are all possible combinations of geographic trend and the four environmental variables (main fertility gradient, slope, pH and C:N ratio individually). The $r^2$ values were averaged across any model containing the variable. The last four columns are the model-averaged standardized regression parameters for the soil and topographic variables in the “all environmental variables” models. Clay indicates the main soil fertility gradient. *** indicates that this soil variable had a significant Bonferroni-corrected p-value in a model with strong support that included this variable.

| Species | Family | Best "Trend/Env/ Rho" Model | $r^2$ | Best "All Env Variables" Model | $r^2$ | Main Fertility Gradient | Slope | pH | C:N Ratio |
|---|---|---|---|---|---|---|---|---|---|
| *Bocageopsis multiflora* (Mart.) R. E. Fries | Annonaceae | trend+ρ | 0.19 | C:N ratio | 0.13 | -0.012 | -0.033 | 0.010 | -0.161 |
| *Duguetia chrysea* Maas | Annonaceae | trend | 0.21 | slope+pH | 0.27 | -0.009 | 0.234 | -0.110 | 0.026 |
| *Duguetia echinophora* R. E. Fries | Annonaceae | env | 0.33 | clay+pH | 0.29 | 0.229*** | -0.009 | 0.094 | -0.083 |
| *Duguetia manausensis* Maas & Boon | Annonaceae | env | 0.24 | clay | 0.15 | 0.187 | -0.120 | 0.085 | 0.009 |
| *Fusaea longifolia* (Aubl.) Saff. | Annonaceae | env | 0.27 | trend+clay+pH | 0.35 | 0.318*** | -0.030 | 0.111 | -0.123 |
| *Unonopsis duckei* R. E. Fries | Annonaceae | trend | 0.16 | pH | 0.17 | 0.128 | -0.030 | 0.306 | 0.027 |
| *Geissospermum argenteum* Woods. | Apocynaceae | env | 0.42 | clay+pH | 0.38 | 0.522*** | -0.077 | 0.228 | 0.031 |
| *Euterpe precatoria* Mart. | Arecaceae | trend | 0.22 | trend | 0.22 | -0.010 | 0.060 | 0.017 | -0.048 |
| *Oenocarpus bacaba* Mart. | Arecaceae | null | NA | C:N ratio | 0.08 | -0.002 | 0.008 | 0.031 | -0.120 |
| *Jacaranda copaia* (Aubl.) D. Don | Bignoniaceae | ρ | 0.21 | clay+pH | 0.26 | 0.098 | 0.016 | 0.267*** | -0.003 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Scleronema micranthum* (Ducke) Ducke | Bombacaceae | null | NA | slope | 0.12 | -0.022 | 0.159 | -0.054 | -0.012 |
| *Protium altsonii* Sandw. | Burseraceae | trend | 0.11 | trend+clay+pH | 0.28 | 0.235 | -0.019 | 0.177 | -0.042 |
| *Protium apiculatum* Swart | Burseraceae | trend | 0.18 | trend | 0.18 | -0.108 | -0.001 | -0.059 | -0.022 |
| *Protium decandrum* (Aubl.) March. | Burseraceae | trend | 0.17 | clay+pH | 0.22 | 0.201 | 0.006 | 0.078 | -0.098 |
| *Protium guianense* (Aubl.) March. | Burseraceae | ρ | 0.20 | trend+clay | 0.35 | 0.289 | 0.036 | -0.062 | 0.003 |
| *Protium hebetatum* Daly | Burseraceae | trend | 0.32 | trend+clay | 0.37 | -0.249 | 0.067 | -0.082 | -0.014 |
| *Protium occultum* Daly | Burseraceae | env | 0.30 | trend+clay+pH | 0.32 | 0.260*** | 0.052 | 0.228 | 0.035 |
| *Protium paniculatum* Engl. | Burseraceae | trend | 0.32 | trend+clay | 0.36 | -0.270*** | 0.031 | -0.106 | 0.060 |
| *Protium strumosum* Daly | Burseraceae | trend | 0.19 | pH | 0.17 | 0.076 | -0.026 | -0.158 | 0.005 |
| *Protium trifoliolatum* Engl. | Burseraceae | trend | 0.23 | trend | 0.23 | 0.035 | -0.019 | -0.079 | 0.045 |
| *Tetragastris panamensis* (Engl.) Kuntze | Burseraceae | env+ρ | 0.47 | clay | 0.37 | -0.349*** | 0.070 | 0.141 | 0.030 |
| *Pourouma bicolor* Mart. | Cecropiaceae | null | NA | null | NA | 0.003 | 0.026 | 0.031 | -0.015 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Pourouma guianensis* Aubl. ssp. guianensis | Cecropiaceae | env | 0.36 | clay+slope+pH | 0.36 | -0.342*** | 0.144 | 0.112 | -0.045 |
| *Pourouma tomentosa* Miq. | Cecropiaceae | null | NA | clay | 0.07 | -0.197 | -0.002 | -0.010 | -0.124 |
| *Goupia glabra* Aubl. | Celastraceae | null | NA | null | NA | 0.009 | -0.023 | 0.022 | -0.044 |
| *Couepia caryophylloides* Benoist | Chryso-balanaceae | env | 0.32 | clay+slope | 0.31 | -0.433*** | -0.241 | -0.010 | 0.015 |
| *Couepia guianensis* Aubl. (Miq.) Prance | Chryso-balanaceae | env+ρ | 0.44 | clay+pH | 0.33 | -0.356*** | 0.031 | -0.152 | 0.062 |
| *Couepia longipendula* Pilger | Chryso-balanaceae | trend | 0.49 | trend+C:N ratio | 0.56 | -0.176 | -0.072 | 0.001 | 0.196 |
| *Hirtella bicornis* Mart. & Zucc. var. pubescens Ducke | Chryso-balanaceae | env | 0.38 | pH | 0.36 | 0.015 | -0.028 | 0.431*** | 0.027 |
| *Hirtella rodriguesii* Prance | Chryso-balanaceae | trend | 0.28 | clay+pH+C:N ratio | 0.37 | 0.166*** | -0.053 | 0.243*** | 0.116 |
| *Licania canescens* Benoist | Chryso-balanaceae | null | NA | clay | 0.10 | 0.129 | -0.097 | -0.003 | -0.044 |
| *Licania caudata* Prance | Chryso-balanaceae | null | NA | null | NA | 0.092 | 0.007 | 0.119 | 0.009 |
| *Licania heteromorpha* Benth. | Chryso-balanaceae | trend | 0.33 | trend+slope | 0.38 | -0.033 | 0.156 | -0.089 | 0.004 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Licania impressa* Prance | Chryso-balanaceae | trend+env | 0.58 | trend+clay+slope+pH | 0.57 | 0.700*** | 0.245 | 0.210 | 0.019 |
| *Licania micrantha* Miq. | Chryso-balanaceae | trend | 0.35 | trend+slope+pH | 0.47 | -0.074 | 0.134 | 0.194 | 0.016 |
| *Licania oblongifolia* Standl. | Chryso-balanaceae | env | 0.28 | slope+C:N ratio | 0.23 | 0.051 | 0.353*** | -0.090 | 0.191 |
| *Buchenavia grandis* Ducke | Combretaceae | null | NA | slope | 0.08 | -0.032 | -0.137 | 0.036 | -0.018 |
| *Sloanea floribunda* Spruce ex Benth. | Elaeocarpaceae | null | NA | null | NA | 0.014 | 0.017 | -0.048 | -0.028 |
| *Conceveiba hostmannii* Benth. | Euphorbiaceae | trend+ρ | 0.62 | trend | 0.59 | 0.002 | 0.003 | 0.033 | 0.006 |
| *Drypetes variabilis* Uittien | Euphorbiaceae | trend+env | 0.36 | trend+clay+pH | 0.34 | 0.314*** | 0.017 | 0.272*** | -0.069 |
| *Hevea guianensis* Aubl. | Euphorbiaceae | null | NA | null | NA | -0.002 | -0.006 | 0.111 | -0.079 |
| *Mabea caudata* Pax & K. Hoffm. | Euphorbiaceae | env | 0.28 | C:N ratio | 0.18 | -0.061 | 0.114 | 0.180 | -0.389*** |
| *Endopleura uchi* (Huber) Cuatr. | Humiriaceae | null | NA | null | NA | 0.062 | -0.015 | -0.024 | 0.009 |
| *Sacoglottis mattogrossensis* Malme var. subintegra (Ducke) | Humiriaceae | ρ | 0.25 | clay+slope | 0.20 | 0.294*** | 0.107 | 0.052 | 0.006 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Vantanea parviflora* Lam. | Humiriaceae | env | 0.28 | clay+C:N ratio | 0.28 | 0.078 | -0.025 | -0.025 | 0.321*** |
| *Lacistema aggregatum* (Berg.) Rusby | Lacistemataceae | null | NA | null | NA | -0.008 | -0.029 | 0.007 | -0.031 |
| *Aniba burchellii/panurensis* s.l. | Lauraceae | null | NA | slope | 0.10 | -0.013 | 0.170 | 0.006 | 0.007 |
| *Licaria cannella* (Meissn.) Kosterm. | Lauraceae | null | NA | pH | 0.13 | 0.055 | -0.021 | -0.232 | 0.002 |
| *Mezilaurus* sp. nov. 1 | Lauraceae | null | NA | pH | 0.06 | 0.079 | -0.059 | 0.189 | -0.007 |
| *Ocotea ceanothifolia* (Nees) Mez | Lauraceae | env | 0.32 | slope+C:N ratio | 0.32 | 0.021 | 0.362*** | 0.016 | -0.122 |
| *Ocotea cinerea* van der Werff | Lauraceae | trend | 0.12 | trend+C:N ratio | 0.21 | -0.003 | -0.005 | -0.079 | 0.119 |
| *Ocotea cujumari* Mart. | Lauraceae | null | NA | clay+pH | 0.21 | 0.228 | -0.015 | -0.101 | 0.001 |
| *Ocotea percurrens* Vicentini | Lauraceae | ρ | 0.26 | trend+clay | 0.33 | 0.122 | 0.005 | -0.007 | -0.099 |
| *Corythophora alta* R. Knuth | Lecythidaceae | trend+ρ | 0.63 | trend+clay | 0.61 | 0.132 | -0.050 | -0.011 | -0.094 |
| *Couratari stellata* A. C. Smith | Lecythidaceae | trend+ρ | 0.53 | trend+slope | 0.49 | -0.017 | -0.113 | -0.043 | -0.020 |
| *Eschweilera atropetiolata* Mori | Lecythidaceae | trend | 0.65 | trend+clay | 0.68 | 0.104 | -0.054 | -0.016 | 0.007 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Eschweilera collina* Eyma | Lecythidaceae | trend+env | 0.46 | trend+clay+slope+C:N ratio | 0.45 | -0.466*** | -0.253 | -0.021 | -0.353 |
| *Eschweilera coriacea* (A. DC.) Mori | Lecythidaceae | env | 0.59 | clay | 0.58 | -1.246*** | -0.015 | 0.017 | -0.037 |
| *Eschweilera cyathiformis* Mori | Lecythidaceae | trend | 0.23 | trend | 0.23 | 0.154 | 0.050 | 0.066 | 0.009 |
| *Eschweilera grandiflora* (Aubl.) Sandw. | Lecythidaceae | null | NA | clay | 0.10 | 0.112 | -0.069 | -0.090 | -0.002 |
| *Eschweilera micrantha* (Berg) Miers | Lecythidaceae | trend+env+ρ | 0.67 | trend+clay+pH +C:N ratio | 0.63 | 0.466*** | 0.013 | -0.454 | 0.234 |
| *Eschweilera pseudodecolorans* Mori | Lecythidaceae | env+ρ | 0.51 | trend+clay+pH | 0.45 | 0.634*** | -0.033 | 0.270 | -0.075 |
| *Eschweilera rankiniae* Mori | Lecythidaceae | trend | 0.40 | trend+slope | 0.49 | -0.020 | -0.172 | -0.117 | 0.042 |
| *Eschweilera romeu-cardosoi* Mori | Lecythidaceae | trend+env | 0.56 | trend+clay | 0.56 | 0.830*** | -0.015 | -0.042 | 0.007 |
| *Eschweilera truncata* A. C. Smith | Lecythidaceae | trend | 0.64 | trend+pH | 0.71 | -0.003 | 0.006 | -0.478 | 0.014 |
| *Eschweilera wachenheimii* (Benoist) Sandw. | Lecythidaceae | null | NA | null | NA | 0.025 | -0.053 | 0.024 | -0.020 |
| *Gustavia elliptica* Mori | Lecythidaceae | trend+ρ | 0.33 | pH | 0.26 | 0.171 | 0.142 | -0.291*** | 0.043 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Lecythis barnebyi* Mori | Lecythidaceae | env | 0.30 | pH+C:Nratio | 0.27 | 0.071 | 0.037 | -0.355*** | -0.235 |
| *Lecythis poiteaui* Berg | Lecythidaceae | ρ | 0.07 | trend+clay+C:N ratio | 0.28 | -0.195 | -0.014 | -0.020 | -0.176 |
| *Bocoa viridiflora* (Ducke) Cowan | Leguminosae | env+ρ | 0.53 | trend+slope+pH | 0.53 | 0.006 | 0.196*** | -0.280*** | -0.040 |
| *Eperua glabriflora* (Ducke) Cowan | Leguminosae | ρ | 0.16 | trend | 0.20 | -0.004 | 0.006 | 0.102 | 0.069 |
| *Inga paraensis* Ducke | Leguminosae | null | NA | pH | 0.08 | -0.023 | -0.024 | -0.100 | -0.016 |
| *Macrolobium limbatum* Benth. | Leguminosae | env | 0.30 | slope+pH | 0.29 | 0.008 | 0.259*** | 0.141 | -0.016 |
| *Paramachaerium ormosioides* (Ducke) Ducke | Leguminosae | trend | 0.34 | trend+clay | 0.44 | 0.401*** | -0.004 | 0.170*** | -0.016 |
| *Sclerolobium* sp. 2 | Leguminosae | trend | 0.18 | trend | 0.18 | -0.060 | -0.006 | -0.013 | 0.012 |
| *Swartzia ingifolia* Ducke | Leguminosae | trend | 0.12 | pH | 0.14 | 0.001 | 0.032 | -0.204 | -0.016 |
| *Swartzia polyphylla* DC. | Leguminosae | null | NA | null | NA | 0.020 | 0.011 | 0.014 | 0.050 |
| *Swartzia recurva* Poepp. | Leguminosae | trend+env | 0.40 | trend+slope+pH | 0.38 | -0.022 | 0.157 | -0.422*** | 0.014 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Swartzia reticulata* Ducke | Leguminosae | env | 0.46 | clay | 0.42 | 0.674*** | -0.082 | 0.048 | -0.037 |
| *Tachigali plumbea* Ducke | Leguminosae | null | NA | null | NA | -0.003 | -0.028 | 0.024 | -0.024 |
| *Zygia racemosa* (Ducke) Barn. & Grimes | Leguminosae | null | NA | null | NA | 0.019 | 0.016 | -0.055 | -0.010 |
| *Mouriri angulicosta/duckeana/duckeanoides* | Melastomataceae | null | NA | slope | 0.11 | 0.007 | -0.171 | 0.098 | -0.001 |
| *Guarea humaitensis* Penn. | Meliaceae | trend+env | 0.43 | trend+clay+pH+C:N ratio | 0.41 | -0.311*** | -0.092 | -0.108 | 0.223*** |
| *Siparuna decipiens* (Tul.) A. DC. | Monimiaceae | null | NA | pH | 0.07 | 0.021 | 0.011 | 0.102 | -0.004 |
| *Brosimum rubescens* Taub. | Moraceae | trend | 0.15 | clay | 0.16 | -0.247*** | 0.008 | -0.110 | -0.048 |
| *Helianthostylis sprucei* Baill. | Moraceae | trend+env | 0.41 | trend+clay+slope+C:N ratio | 0.41 | 0.285*** | 0.080 | 0.006 | 0.420*** |
| *Helicostylis tomentosa* (Planch. & Endl.) Rusby | Moraceae | trend | 0.24 | trend | 0.24 | -0.019 | 0.012 | 0.003 | -0.043 |
| *Maquira sclerophylla* (Ducke) C. C. Berg | Moraceae | env | 0.28 | clay+pH | 0.25 | 0.275*** | -0.062 | 0.116 | -0.050 |
| *Naucleopsis caloneura* (Huber) Ducke | Moraceae | trend | 0.11 | pH | 0.06 | -0.023 | 0.027 | -0.084 | 0.064 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Pseudolmedia laevis* (Ruiz & Pav.) Macbr. | Moraceae | env+ρ | 0.33 | clay | 0.08 | -0.111 | 0.013 | 0.011 | 0.004 |
| *Trymatococcus amazonicus* Poepp. & Endl. | Moraceae | trend | 0.12 | clay | 0.12 | -0.148 | 0.024 | -0.033 | 0.068 |
| *Iryanthera dialyandra* Warb. | Myristacaceae | null | NA | null | NA | 0.032 | 0.019 | 0.009 | -0.042 |
| *Iryanthera juruensis* Warb. | Myristacaceae | env | 0.26 | clay+slope | 0.25 | -0.127 | 0.228*** | 0.013 | 0.023 |
| *Virola calophylla* Warb. | Myristacaceae | null | NA | slope | 0.13 | 0.004 | 0.212 | -0.024 | 0.054 |
| *Virola sebifera* Aubl. | Myristacaceae | env+ρ | 0.38 | C:N ratio | 0.22 | -0.013 | 0.022 | -0.069 | -0.373*** |
| *Virola venosa* (Benth.) Warb. | Myristacaceae | null | NA | null | NA | 0.006 | -0.045 | 0.018 | 0.017 |
| *Neea* sp. 2 | Nyctaginaceae | null | NA | null | NA | -0.003 | -0.024 | -0.027 | 0.024 |
| *Heisteria laxiflora* Engl. | Olacaceae | trend | 0.11 | trend+pH | 0.24 | -0.042 | 0.050 | -0.308 | 0.047 |
| *Minquartia guianensis* Aubl. | Olacaceae | null | NA | C:N ratio | 0.06 | 0.066 | 0.040 | 0.072 | -0.075 |
| *Chrysophyllum sanguinolentum* (Pierre) Baehni | Sapotaceae | env | 0.47 | clay+slope+pH | 0.47 | -0.452*** | 0.177 | -0.109 | -0.006 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Ecclinusa guianensis* Eyma | Sapotaceae | null | NA | trend+clay+pH | 0.23 | 0.131 | -0.009 | 0.110 | 0.013 |
| *Manilkara bidentata* (A. DC.) Chev. | Sapotaceae | null | NA | null | NA | 0.005 | 0.011 | 0.016 | -0.008 |
| *Manilkara huberi* (Ducke) Chev. | Sapotaceae | trend+env | 0.59 | trend+clay+slope+pH | 0.57 | 0.445*** | -0.119 | 0.373*** | -0.050 |
| *Micropholis casiquiarensis/mensalis* | Sapotaceae | null | NA | clay | 0.08 | -0.094 | 0.012 | 0.017 | -0.018 |
| *Micropholis guyanensis* (A. DC.) Pierre | Sapotaceae | env | 0.29 | clay | 0.26 | -0.445*** | -0.002 | -0.024 | 0.040 |
| *Micropholis venulosa* (Mart. & Eichl.) Pierre | Sapotaceae | null | NA | null | NA | -0.051 | -0.003 | 0.059 | 0.021 |
| *Micropholis williamii* Aubr. & Pellegr. | Sapotaceae | trend | 0.16 | trend | 0.16 | 0.001 | 0.032 | -0.071 | 0.053 |
| *Pouteria ambelaniifolia* (Sandw.) Penn. | Sapotaceae | ρ | 0.11 | pH | 0.06 | -0.010 | -0.003 | -0.063 | 0.006 |
| *Pouteria anomala* (Pires) Penn. | Sapotaceae | env | 0.54 | slope+pH | 0.52 | -0.075 | 0.470*** | -0.820*** | 0.075 |
| *Pouteria caimito* (R. & P.) Radlk. | Sapotaceae | trend | 0.19 | trend | 0.19 | -0.036 | 0.035 | 0.015 | 0.005 |
| *Pouteria campanulata* Baehni | Sapotaceae | null | NA | pH | 0.06 | -0.052 | 0.013 | -0.065 | 0.007 |
| *Pouteria cladantha* Sandw. | Sapotaceae | trend | 0.20 | clay+pH | 0.22 | 0.110 | -0.067 | 0.098 | -0.048 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Pouteria cuspidata* (A. DC.) Baehni | Sapotaceae | null | NA | slope | 0.08 | -0.023 | 0.137 | 0.057 | 0.061 |
| *Pouteria engleri* Eyma | Sapotaceae | trend | 0.14 | slope+pH | 0.21 | 0.071 | -0.137 | 0.210 | 0.040 |
| *Pouteria ericoides* Penn. | Sapotaceae | null | NA | null | NA | 0.068 | 0.006 | 0.020 | -0.005 |
| *Pouteria erythrochrysa* Penn. | Sapotaceae | trend | 0.20 | trend | 0.20 | 0.046 | 0.005 | -0.026 | 0.006 |
| *Pouteria eugeniifolia* (Pierre) Baehni | Sapotaceae | trend | 0.18 | trend+slope+pH | 0.32 | 0.014 | 0.133 | 0.140 | 0.009 |
| *Pouteria filipes* Eyma | Sapotaceae | null | NA | pH | 0.09 | -0.019 | 0.033 | -0.131 | 0.027 |
| *Pouteria freitasii* Penn. | Sapotaceae | trend | 0.20 | trend+pH | 0.26 | 0.027 | 0.023 | 0.137 | 0.019 |
| *Pouteria guianensis* Aubl. | Sapotaceae | trend+env | 0.39 | trend+slope+pH | 0.36 | 0.007 | 0.229*** | 0.294*** | -0.032 |
| *Pouteria hispida* Eyma | Sapotaceae | null | NA | slope | 0.05 | 0.016 | -0.066 | -0.001 | 0.027 |
| *Pouteria macrophylla* (Lam.) Eyma | Sapotaceae | env+ρ | 0.49 | clay+pH | 0.39 | 0.353*** | -0.079 | 0.139 | 0.011 |
| *Pouteria minima* Penn. | Sapotaceae | trend | 0.23 | clay+pH | 0.28 | -0.143 | 0.006 | -0.136 | 0.109 |
| *Pouteria reticulata* (Engl.) Eyma | Sapotaceae | trend | 0.29 | pH+C:Nratio | 0.34 | -0.019 | 0.004 | -0.202*** | 0.107 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Pouteria retinervis* Penn. | Sapotaceae | trend | 0.12 | C:N ratio | 0.13 | 0.008 | -0.009 | -0.003 | 0.184 |
| *Pouteria rostrata* (Huber) Baehni | Sapotaceae | env | 0.30 | pH | 0.29 | -0.013 | 0.031 | -0.410*** | 0.005 |
| *Pouteria venosa* (Mart.) Baehni | Sapotaceae | null | NA | null | NA | -0.013 | -0.032 | 0.043 | -0.041 |
| *Pouteria vernicosa* Penn. | Sapotaceae | trend | 0.24 | trend | 0.24 | -0.036 | 0.054 | -0.080 | 0.118 |
| *Sterculia parviflora* (Ducke) Taylor | Sterculiaceae | trend | 0.12 | trend+C:N ratio | 0.20 | -0.017 | 0.007 | 0.042 | -0.121 |
| *Theobroma sylvestre* Mart. | Sterculiaceae | trend | 0.17 | trend | 0.17 | 0.040 | 0.003 | -0.039 | 0.011 |
| *Amphirrhox surinamensis* Eichl. | Violaceae | null | NA | pH | 0.06 | 0.076 | -0.009 | 0.137 | 0.020 |
| *Rinorea guianensis* Aubl. var. subintegrifolia | Violaceae | env | 0.26 | clay+slope | 0.25 | 0.443*** | 0.203 | 0.037 | -0.017 |
| *Rinorea racemosa* (Mart.) Kuntze | Violaceae | null | NA | slope | 0.06 | 0.010 | -0.087 | -0.023 | -0.014 |
| *Erisma bicolor* Ducke | Vochysiaceae | trend | 0.13 | trend+slope | 0.20 | 0.057 | -0.093 | -0.007 | 0.036 |
| *Qualea labouriauna* Paule | Vochysiaceae | env | 0.30 | clay+pH | 0.28 | 0.198 | -0.048 | 0.212 | -0.043 |

Appendix 6. Results from ordinary least squares (OLS) models with all possible combinations of one geographic trend variable and four environmental variables (main fertility gradient, slope, pH and C:N ratio individually) as explanatory variables. Because $PO_4$ and available water were the least important variables in the community level analysis, they were not included in the species level analyses. The last three rows contain the sum of $AIC_c$ weights and the sum of species with the best or strongly-supported model for any model containing the variable listed.

| Model | $AIC_c$ weights | Best model<br># species | Best model or strong support<br># species |
|---|---|---|---|
| Null | 5.1% | 18 (13%) | 58 (43%) |
| pH | 5.0% | 15 (11%) | 63 (47%) |
| Trend | 5.3% | 13 (10%) | 82 (61%) |
| Main fertility gradient | 5.4% | 12 (9%) | 75 (56%) |
| Main fertility gradient+pH | 5.7% | 11 (8%) | 76 (56%) |
| Slope | 3.8% | 8 (6%) | 54 (40%) |
| Trend+main fertility gradient | 4.5% | 8 (6%) | 67 (50%) |
| C:N ratio | 3.2% | 6 (4%) | 50 (37%) |
| Trend+main fertility gradient+pH | 3.9% | 6 (4%) | 46 (34%) |
| Trend+slope+pH | 2.6% | 5 (4%) | 30 (22%) |
| Main fertility gradient+slope | 3.5% | 4 (3%) | 57 (42%) |
| Slope+pH | 4.0% | 4 (3%) | 64 (47%) |
| Trend+slope | 3.4% | 4 (3%) | 50 (37%) |
| Trend+C:N ratio | 3.1% | 3 (2%) | 54 (40%) |
| trend+pH | 3.4% | 3 (2%) | 59 (44%) |
| Main fertility gradient+slope+pH | 3.5% | 2 (1%) | 55 (41%) |
| pH+C:N ratio | 2.8% | 2 (1%) | 45 (33%) |
| Slope+C:N ratio | 2.6% | 2 (1%) | 47 (35%) |
| Trend+main fertility gradient+pH+C:N ratio | 2.0% | 2 (1%) | 31 (23%) |
| Trend+main fertility gradient+slope+C:N ratio | 1.8% | 2 (1%) | 16 (12%) |
| trend+main fertility gradient+slope+pH | 2.3% | 2 (1%) | 34 (25%) |
| Main fertility gradient+C:N ratio | 3.3% | 1 (1%) | 60 (44%) |
| Main fertility gradient+pH+C:N ratio | 3.1% | 1 (1%) | 54 (40%) |
| Trend+main fertility gradient+C:N ratio | 2.6% | 1 (1%) | 37 (27%) |
| Main fertility gradient+slope+C:N ratio | 2.0% | 0 (0%) | 36 (27%) |
| Main fertility gradient+slope+pH+C:N ratio | 1.7% | 0 (0%) | 27 (20%) |
| Slope+pH+C:N ratio | 2.1% | 0 (0%) | 37 (27%) |
| Trend+main fertility gradient+slope | 2.6% | 0 (0%) | 35 (26%) |
| Trend+main gradient+slope+pH+C:N ratio | 1.2% | 0 (0%) | 10 (7%) |
| Trend+pH+C:N ratio | 1.7% | 0 (0%) | 31 (23%) |
| Trend+slope+C:N ratio | 1.7% | 0 (0%) | 30 (22%) |
| Trend+slope+pH+C:N ratio | 1.2% | 0 (0%) | 15 (11%) |

Appendix 7. Comparison of the percentage of individual species with significant relationships with soil and topographic variables from different tropical forests. For studies with “trees” listed as the species type, no subset of species was based on the abundance of species rather than on a specific taxon. Total area sampled was calculated by adding up the size of the individual plots. The statistical methods used for determining whether a species had a significant relationship with a soil variable or habitat varied among studies.

uAppendix 7

| Authors | Year | Country | No. of spp. | % species with significant soil relationships | Species type | Soil variables or habitats | Study area size (ha) | No. of plots | Total area sampled (ha) |
|---|---|---|---|---|---|---|---|---|---|
| Hall et al. | 2004 | Central African Republic | 4 | 75 | *Entandrophragma* | chemical, physical, topographic variables | 100 | 1 | 100 |
| Sri-ngernyuang et al. | 2003 | Thailand | 17 | 65 | Lauraceae | topographic variables | 7.5 | 1 | 7.5 |
| Paoli et al. | 2006 | Borneo | 22 | 82 | Dipterocarpaceae | chemical, physical variables | 340 | 30 | 4.8 |
| Svenning | 1999 | Ecuador | 31 | 65 | palms | topographic-, hydrologic-, other-based habitats | 50 | 118 | 4.72 |
| Phillips et al. | 2003 | Peru | 34 | 77 | trees | pleistocene, holocene substrate habitats | 1000000 | 88 | 8.8 |
| Webb & Peart | 2000 | Borneo | 49 | 35 | trees | topographic-based habitats | 150 | 28 | 4.48 |
| Miyamoto et al. | 2003 | Borneo | 55 | 80 | trees | elevation, humus depth | 200 | 2 | 2 |
| Cannon & Leighton | 2004 | Borneo | 55 | 67 | trees | substrate, elevation, humus depth-based habitats | 1500 | 69 | 6.9 |

Appendix 7 (con't)

| Authors | Year | Country | No. of spp. | % species with significant soil relationships | Species type | Soil variables or habitats | Study area size (ha) | No. of plots | Total area sampled (ha) |
|---|---|---|---|---|---|---|---|---|---|
| Clark et al. | 1999 | Costa Rica | 132 | 25 | trees | hydrologic-, substrate- and topographic-based habitats | 573 | 110 | 1.1 |
| Clark et al. | 1998 | Costa Rica | 9 | 78 | trees | soil type, topographic position, slope | 216 | NA | NA |
| Tuomisto et al. | 2002 | Ecuador | 143 | 47 | melostomes and pteridophytes | chemical variables | 50000 | 27 | 6.75 |
| Harms et al. | 2001 | Panama | 171 | 51 | trees | topographic-, hydrologic- habitats | 50 | 1 | 50 |
| Yamada et al. | 2006 | Borneo | 10 | 80 | trees | topographic-, soil texture- habitats | 52 | 1 | 52 |
| Bohlman et al. | this study | Brazil | 135 | 30 | trees | chemical, physical, topographic variables | 100000 | 66 | 66 |