DTU Library

Relationships between plant species richness and soil pH at the level of biome and ecoregion in Brazil

Crespo Mendes, Natalia; Laurent, Alexis; Bruun, Hans Henrik; Hauschild, Michael Zwicky

Relationships between plant species richness and soil pH at the level of biome and ecoregion in Brazil

Natalia Crespo-Mendes[1*], Alexis Laurent[1], Hans Henrik Bruun[2] and Michael Zwicky Hauschild[1]

[1] Division for Quantitative Sustainability Assessment (QSA), Department of Management Engineering, Technical University of Denmark (DTU), 2800 Kgs. Lyngby, Denmark

[2] Department of Biology, University of Copenhagen (KU), 2100 Copenhagen Ø, Denmark

[*] To whom correspondence should be addressed; e-mail: ncme@dtu.dk

Abstract

Soil pH has been used to indicate how changes in soil acidity can influence species loss. The correlation between soil pH and plant species richness has mainly been studied in North America and Europe, while there is a lack of studies exploring Tropical floras. Here, our aim was therefore to investigate the relationships between terrestrial plant species richness and soil pH for the large Brazilian flora, with spatial differentiation into biomes and ecoregions. Data of plant species occurrences and soil pH in Brazil were compiled from public databases into a geo-referenced inventory of 29712 terrestrial plants species with a harmonized nomenclature. Based on the pH range, over which each species had been observed, the species richness for each unit of soil pH was determined and plotted as a function of pH for the 6 biomes and 47 ecoregions of Brazil. Lognormal distributions were found for entire Brazil ($R^2 = 0.999$), the six biomes ($R^2 > 0.955$) and for 40 out of 45 ecoregions, for which a sufficient number of observations was available (R^2 of 0.830-1.000). Similar distribution patterns were observed when limiting the study scope to range-restricted species, i.e. species only occurring in a single ecoregion in Brazil. Species richness is an indicator of plant biodiversity and we recommend a combined use of species richness for all species and for range-restricted species to address the overall status of the terrestrial plant ecosystem as well as the potential loss of unique species within it, including endemic species. We additionally propose that the developed inventory and the observed sensitivity distributions serve as basis for life cycle impact assessment of terrestrial acidification.

Keywords: Brazilian flora, Biodiversity conservation, Range-restricted species, Soil acidity, Ecoregion, Biome.

1. INTRODUCTION

Biodiversity loss is a worldwide concern and the central point of studies defining global conservation priorities (Myers et al., 2000; Orme et al., 2005; Brooks et al., 2006). Habitat loss and habitat change due to anthropogenic pressures are the prime drivers of biodiversity loss. To identify patterns of species loss in support of conservation policies, biodiversity indicator results can be analyzed together with physicochemical properties of the environment.

With regard to plants, the influences of soil acidity and availability of soil nutrients on plant species richness have received considerable attention in the scientific literature. The majority of studies have addressed North American and European regions (Gough et al., 2000; Roem and Berendse, 2000; Pärtel, 2002; Crawley et al., 2005; Duprè et al., 2010; Stevens et al., 2010), with few exceptions such as the study by Azevedo et al. (2013) focusing on different world biomes. Most studies found a correlation between soil pH and plant species richness, but they also indicated other possible drivers of change in the species richness, such as precipitation (Gentry, 1988), latitude (Duprè et al., 2010) and nitrogen deposition, which often accompanies airborne acidification (Duprè et al., 2010; Stevens et al., 2010). In this context, analyses considering physicochemical properties and biodiversity indicators together, such as soil pH and species richness respectively, contribute to identify specificities within ecosystems. In Pärtel (2002), for example, analyses considering where the pools of species are suited, whether for low or high pH soil, may relate to evolutionary history on local-scale diversity patterns. Additionally, Azevedo et al. (2013) use the relationships between species richness and soil pH to assess the potential effects that acidifying substances might cause in terrestrial ecosystems.

Furthermore, analyzing the relationships between species richness and soil pH within regions with common climate, vegetation, geology, etc. may help capture regional differences that have not been identified in studies covering more extensive territorial divisions. Ecoregions, which are biogeographic units containing a distinct assemblage of natural communities sharing a large proportion of species, dynamics, and environmental conditions (Olson et al., 2001), appear as a good choice for this type of investigation. Moreover, besides the concern with the loss of wide-ranging species, knowing what unique – or range-restricted - species of each ecoregion are can bring the advantage of an analysis focused on the preservation of species that could potentially be extinct, preventing the loss of biodiversity.

Brazil is the country in the world hosting the largest floristic diversity, with more than 30000 species of higher plants recorded (Forzza et al., 2012; Brazil Flora G, 2015). However, there is a lack of studies exploring the relationships between Tropical plant species richness and soil pH. The Brazilian Flora Checklist (Brazil Flora G, 2015) provides information on the distribution of species into biomes and estimates the proportion of endemic species to Brazil (Zappi et al., 2015; Costa and Peralta, 2015; Prado et al., 2015). Nevertheless, information on species richness and range size of

species, as well as the relationships between species richness and soil pH are not provided at the finer resolution of ecoregions. The most comprehensive study known to the authors is a global study that observed occurrences of 2409 plant species categorized in 13 terrestrial biomes across the world (Azevedo et al., 2013). The modest number of species covered at global scale, compared to the more than 30000 plant species reported for Brazil alone (Forzza et al., 2012; Brazil Flora G, 2015), renders the representativeness of this and other less ambitious studies questionable, pointing to the relevance of a more comprehensive analysis.

Rather than further analyze patterns of plant diversity in Brazil and the effects of human interactions on the ecosystems, in this study, we aim to 1) investigate the relationships between terrestrial plant species richness and soil pH in Brazil at the level of country, biome and ecoregion, using a large tropical flora, and 2) assess potential differences in these relationships between species, which are unique to single ecoregions, termed "range-restricted species" in the study, and those, which have extended ranges of occurrence.

2. METHODOLOGY

A large methodological challenge is to obtain sufficiently comprehensive and representative data to analyze the relationships between species richness and soil pH. To overcome this challenge, instead of retrieving data through a literature review, which is the approach used in previous studies (Pärtel, 2002; Azevedo et al., 2013), we have compiled information from separate databases of plant species occurrences and soil pH in Brazil, and analyzed the information using the Geographic Information System (GIS) software ArcGIS 10.3.1 (https://www.arcgis.com) and statistical tools. Details of the applied methodology are given in the following sections.

2.1. Occurrences of plant species in Brazil

An occurrence of a plant species refers to a recorded observation of the plant species at a specific location. Only records with information on the genus and species names of the plant, and the latitude and longitude of the observation site were used. The species richness of a region (ecoregion or biome) was defined as the number of different species observed in the region. The generated inventory presents an overview of the occurrence of species over the years and all occurrence data are treated the same, irrespective of the year of observation. Thus, ecological successions are not differentiated in this study, whether they are caused by natural forces or human interactions, such as forest fires or agricultural settlement. Such causality would however be worth investigating in further research work.

2.1.1. Data sources for species occurrence

Data for the Brazilian inventory were extracted from the Global Biodiversity Information Facility (GBIF, 2015; GBIF.org, 2015a,b,c,d,e). GBIF is integrated with the Brazilian platform (Sistema de Informação sobre a Biodiversidade Brasileira - SiBBr) and data quality is ensured by more than 90 endorsed data publishers for Brazil along with additional checks performed by GBIF (GBIF secretariat, 2017). However, the virtue of GBIF data is their exuberance, rather than the accuracy of the individual data entry (see Section 2.4). Collecting data from the GBIF database is believed to ensure that the greatest number of digitized records publicly available is considered in the present study. More details on the GBIF database are provided in Electronic Supplementary Material 1 (ESM-1), Supporting Methods A.

2.1.2. Development of terrestrial plant species inventory

In the extraction of species occurrence data, only records belonging to the kingdom *Plantae* were considered. ArcGIS 10.3.1 was used for processing the extracted data into a georeferenced inventory, and the World Geodetic System 1984 (WGS84) was adopted as the geographic coordinate system. The map of Brazil provided by the Brazilian Institute of Geography and Statistics (IBGE), which is responsible for statistical, geographic, cartographic, geodetic and environmental information in Brazil, was used as a reference for defining the Brazilian territory (IBGE, 2015).

To only use accepted names and avoid double counting of species, a taxonomic alignment was performed and supported by the use of the Taxonomic Name Resolution Service v4.0 (TNRS; Boyle et al., 2013). TNRS is an online application for automated standardization of plant scientific names with reference to existing high-quality taxonomy sources (Boyle et al., 2013). The following taxonomic data sources were used in this study: Missouri Botanical Garden's Tropicos database (i.e. Tropicos, 2015), The Global Compositae Checklist (GCC; Flann, 2015), The Plant List (TPL, 2015) and The International Legume Database and Information Service (ILDIS, 2015). The list of species names was submitted to the TNRS and the configurations set for the analysis are presented in Supporting Methods. Inconsistent records were flagged from the run through TNRS tool and adjustments were performed – see details in the Supporting Methods B (ESM-1). The species habitat inventory list provided by the Brazilian Plant Checklist (Brazil Flora G, 2015) was also integrated into the inventory, enabling exclusion of non-terrestrial plant species.

2.1.3. Spatial resolution

Species in the resulting inventory of terrestrial plants in Brazil were grouped at different spatial resolutions. Several approaches exist for classification of biomes (Olson et al., 2001; Hoekstra et

al., 2005; IBGE, 2004; Ellis and Ramankutty, 2008) and ecoregions in Brazil (Dinerstein et al., 1995; Olson and Dinerstein, 1998; Olson et al., 2001). The current study uses the widely-applied classification delineated by Olson et al. (2001) because it covers both biomes and ecoregions for the entire world and thus provides compatibility with global data sets that may be developed in the future. According to this classification the terrestrial world is subdivided into 14 biomes and 867 ecoregions to better reflect the distribution of the Earth's natural communities and species. The three spatial scales adopted in this study therefore are the ecoregion level (47 ecoregions), the biome level (six biomes) and the whole-Brazil level.

2.1.4. Range-restricted species

In this study, we adopted a Brazilian perspective and defined range-restricted species at the level of ecoregions, i.e. species only occurring in one of Brazil's ecoregions. Species that are range-restricted from this definition may still occur in other locations outside Brazil, but given the size of the country and the high species richness, this classification still provides useful information about the vulnerability to loss of the species. Therefore, the range-restricted species will also include the Brazilian endemic species (which do not occur anywhere else than in Brazil).

2.1.5. Correlation between total species richness and range-restricted species richness

Species richness for all species and for range-restricted species alone were considered as potential indicators of biodiversity in this study. The latter was tested as a biodiversity indicator by evaluating the fraction of range-restricted species (FRS) out of the total species in each of the ecoregions and biomes in Brazil, and by assessing variations across the regions.

2.2. Soil pH data

Data on soil pH were accessed through SoilGrids1km, which is an automated system for global soil mapping and is part of the Global Soil Information Facilities (GSIF), a platform developed by the International Soil Reference and Information Centre (ISRIC) for collating and predicting soil properties and soil classes in 3D at 1km resolution (Hengl et al., 2014). The soil pH maps were generated from point observations and covariation layers, and for Brazil, the Brazilian national soil profile database was used as the main source of points of observation (Hengl et al., 2014). The Brazilian national soil profile database consists of 5086 profiles with a total of 10034 soil horizons (i.e. distinct layers running parallel to the surface) with information on 31 variables covering soil morphological, physical and chemical attributes, one of them being the soil pH (Cooper et al., 2005).

The collected data were processed using ArcGIS v.10.3.1. Brazil was divided into approximately 10.2 million grid cells of 1km x 1km and a soil pH value was extracted for each grid cell. Soil horizons representing a depth range of 0-60 cm were assumed relevant for plant exposure to acidifying or alkalizing substances. To obtain a representative pH value, an arithmetic mean of the average proton concentrations was calculated across this depth range for each grid cell. In the use of these pH data, inconsistencies present in the database were found and addressed. Details are provided in Supporting Methods C (ESM-1).

2.3. Processing of species richness

2.3.1. Species richness distribution

The georeferenced locations for occurrences of plant species (Section 2.1) were matched with the soil pH 1-km^2 grid cells (Section 2.2) to create an inventory of occurrences of plant species at different soil pH values within each of the ecoregions and biomes. From this information, each species could be attributed a range of soil pH, delimited by the lowest and highest pH values at which it has been reported within a given ecoregion or biome. A species may thus not necessarily have been recorded at all intermediary pH values within its range. It may also exist outside the defined pH range even though it has not been recorded. It was however not possible to check the latter, and considering the very high number of observations reported for most ecoregions, it was assumed that the minimum and maximum pH values defining the range are representative of the occurrence of the species. Based on the observed ranges, the species richness distributions as functions of the soil pH were determined at the biome level and at the ecoregion level for both the entire list of species and the list of range-restricted species.

2.3.2. Correlation between species richness and soil pH

Regression analyses were performed to analyze the relationships between soil pH and species richness as determined in Section 2.3.1. Parametric fitting models such as logistic and lognormal distributions are commonly used for species richness distributions (Guisan et al., 2002; Longino et al., 2002; Volkov et al., 2003; McGill et al., 2007; Azevedo et al., 2013; Colwell and Coddington, 1994). Both logistic and lognormal distribution models were tested to identify the one that statistically provided the best fit.

2.4. Uncertainties

Four factors were identified as possible sources of uncertainty in this study: (i) inaccurate georeferencing of occurrences; (ii) errors in taxonomic identification; (iii) classification and/or selection of plant species with terrestrial habitat; and (iv) estimates of soil pH values.

Georeferencing errors in records used in databases such as GBIF can be associated with the lack of accuracy when digitizing the recorded samples from old collections, for which the geographical coordinates are not available or are not easily readable (Maldonado et al., 2015). Inaccurate georeferencing may lead to a species erroneously recorded in the neighboring ecoregion. For these cases, we assumed that the boundaries between ecoregions can change gradually due to the difficulty in limiting the transition zones between the different types of vegetation. Thus, even if a species is counted in the neighboring ecoregion, considering that the transition zones might have similar environmental characteristics and considering the high number of species occurrences for most ecoregions, overall the resulting error is very likely negligible.

Adopting the correct species name is often associated to cases of taxonomic disagreement regarding species delimitations, synonymisation and nomenclatural problems, making such an identification not trivial (Maldonado et al., 2015). The taxonomic alignment step was performed to minimize errors when selecting the accepted species name and its consistency relies on the set of high-quality taxonomy sources (TPL, Tropicos, GCC and ILDIS). The authority of the names and the list of accepted names and synonyms may vary according to the order of consultation of the taxonomy sources. Given the large data set and the large number of species, this source of uncertainty is likely to be negligible.

Regarding the selection of terrestrial plant species, 25223 out of 33166 identified species have the habitat classification available and approximately 86% of these species have explicitly been associated with a terrestrial habitat classification (i.e. 21738 out of the 25223 species root in soil). The uncertainties come from the remaining species for which the habitat was not reported (7943 species out of 33166). Due to the high percentage of terrestrial species in a land-based environment as Brazil, it was assumed that all species with unknown habitat were terrestrial to avoid arbitrary choices on which species to be disregarded as a non-terrestrial among the set of species with unknown habitat. Species with known habitat that do not root in soil, such as aquatic, epiphytic, parasitic and saprophytic species, were disregarded as non-terrestrial species.

The uncertainties related to the estimates of soil pH values come from the statistical modelling of the distribution of soil pH values. The pH values are modelled based on extrapolation from a limited number of measured values. The predictions for each layer are based on the predicted value (mean) and the 90% prediction interval, which can be used to propagate uncertainties in models where soil property maps are used. In this study, only the mean value was used. The width of the

prediction interval varies from 1.5 to 1.6 pH units for all regions and is not expected to impact the observed relationships between species richness and soil pH since there is the same probability of the correct value being higher or lower than the average. Specifically for the pH values the uncertainty about the data can also come from laboratory errors, but this is not considered as a source of uncertainty in the prediction models.

3. RESULTS AND DISCUSSION

3.1. Inventory of terrestrial plant species in Brazil

3.1.1. Inventory for entire Brazil

Disregarding habitat type, this study initially addressed 976345 occurrences of plants in Brazil, representing 33166 different species of Angiosperms, Bryophytes, Gymnosperms, Ferns and Lycophytes. For the same groups of plants, a total of 35639 species have been catalogued in the latest published update of the Brazilian List system (Zappi et al., 2015; Costa and Peralta, 2015; Prado et al., 2015), thus suggesting a high representativeness in the developed inventory. A comparison between the content of the present dataset and the information provided by the Brazilian Flora Checklist (Brazil Flora G, 2015) shows that the Brazilian Flora Checklist has catalogued 97% of the occurrences of plants, which represent 79% of the reported species, addressed in the present inventory. The differences between the two lists may be explained by choices related to the taxonomic alignment, i.e. the authority given to each of the considered taxonomy sources and the different levels of updating (see Section 2.1.2).

After the habitat selection, the inventory of plant species in Brazil with terrestrial habitat comprises 891313 occurrences of plants covering 29712 species. Among these, a total of 8242 plant species were identified as range-restricted (see definition in Section 2.1.4), corresponding to 28% of the total number of plant species in the country. The inventory of terrestrial plant species at country, biome and ecoregion levels, with differentiation of range-restricted species, is given in the Electronic Supplementary Material 2 (ESM-2; Excel file). To summarize the data associated with the inventory, Table S1 in ESM-1 presents an overview of the species number for each biome and ecoregion of Brazil.

3.1.2. Biome level

The plant species counts for the six biomes in Brazil (Table S1, ESM-1) show that the biome Tropical and subtropical moist broadleaf forests is the most species-rich biome with 25774 species, representing 87% of the total number of terrestrial plant species registered in Brazil (see Figure 1a, in dark red). It is also the biome with the highest number of range-restricted species, with 5373

species, representing 65% of the total number of range-restricted species identified in this study (see Figure 1c). The high number of species present in this single biome is consistent with the literature (Costa and Peralta, 2015; Prado et al., 2015 and Zappi et al., 2015). This biome indeed covers the region of the Amazon forest – a known source of great biodiversity in the world located in the North Brazil – and the Southeastern Brazil that concentrates most of the species according to the Brazilian Flora Checklist (Brazil Flora G, 2014). While the Amazon Rainforest is reported to have the highest number of species for Gymnosperms, the Southeastern Brazil has been identified as hosting the highest number of species and endemic species for Angiosperms, Bryophytes, Ferns and Lycophytes (Costa and Peralta, 2015; Prado et al., 2015 and Zappi et al., 2015).
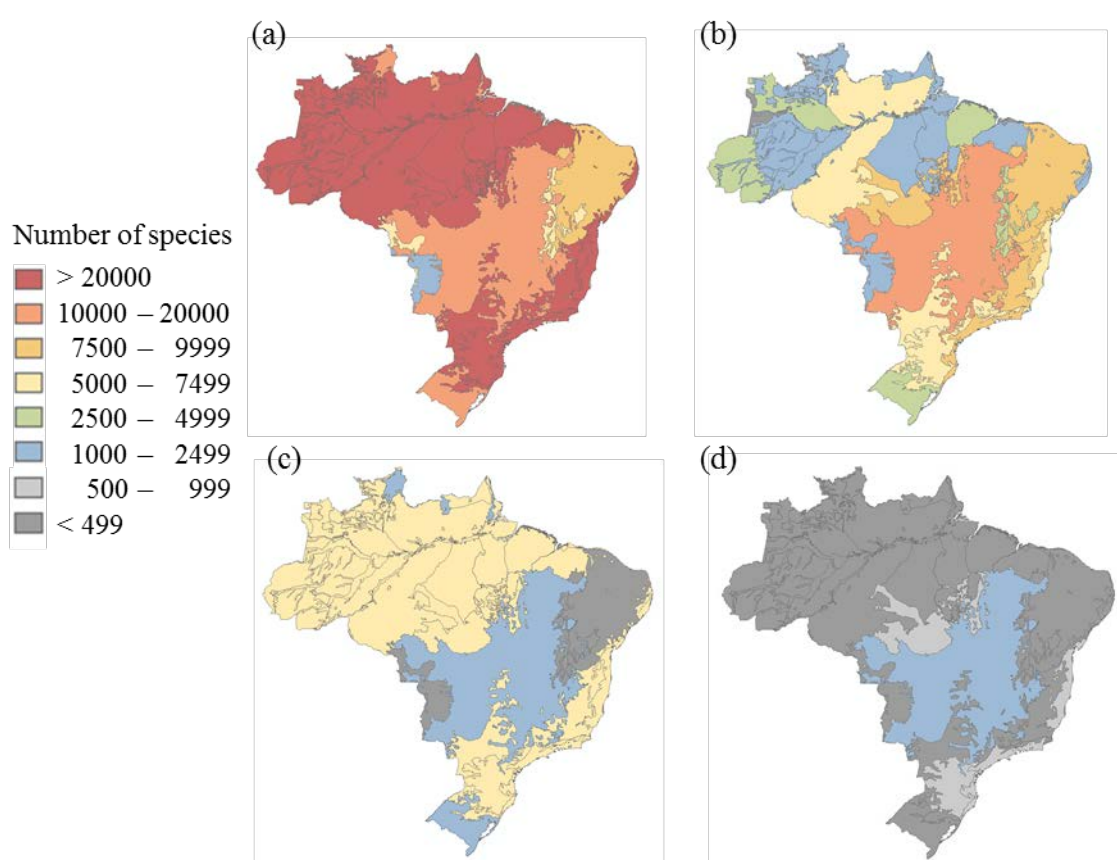

Figure 1. Terrestrial plant species in Brazil: (a) Total species per biome (six biomes), (b) Total species per ecoregion (47 ecoregions), (c) Range-restricted species per biome (six biomes) and (d) Range-restricted species per ecoregion (47ecoregions).

3.1.3. Ecoregion level

At ecoregion level, the plant species counts indicate that out of the 47 ecoregions, the ecoregion Cerrado presents the highest number of terrestrial plant species in Brazil, with 12751 species, thus capturing ca. 43% of all terrestrial species in Brazil and 54% of the species within the biome

Tropical and subtropical grasslands, savannas and shrublands, to which it belongs (see Figure 1b). Besides being the ecoregion with the highest number of species in Brazil, Cerrado presents the highest number of range-restricted species, with 1573 species corresponding to 19% of the total number of range-restricted species identified in Brazil (see Table S1, ESM-1). The detailed information for the remaining ecoregions can be found in Table S1 in ESM-1.

3.1.4. Influence of data collection distribution on the species count

The larger the area, the larger the number of collection spots, and the higher the number of species (species richness) in an ecoregion or biome (see Figures S1-S2 and Table S2, ESM-1). This apparent bias is reduced when the number of collection spots is expressed relative to the total area of the ecoregion or biome. The biome Tropical and subtropical moist broadleaf forests (Figure 1a, in red) is a notable example on how the non-balanced geographic distribution of data collection influences the count of species. Despite having the highest number of species and the highest number of collection spots, the biome Tropical and subtropical moist broadleaf forests presents only 0.01 spot/km^2 (based on the data given in Table S1, ESM-1). Most of the ecoregions, which compose this biome and are located in the North Brazil (Amazon rainforest), have a low number of collection spots per area (< 0.01 spots/km^2), while the number is higher for ecoregions located on the coast or in the southeast of Brazil, as Serra do Mar coastal forests (0.12 spots/km^2) and Bahia coastal forests (0.09 spots/km^2).

The comparison between the distribution of collection spots per biome and ecoregion shows that the concentration of collection spots is higher in the central and southeastern areas of Brazil (see Figure 2). This high concentration of collection spots is not necessarily related to high numbers of species, but rather to the high number of groups of experts and research institutions located in these regions, while access to the Amazon forest regions is limited. Thus, even if the Southeastern region is the largest contributor in numbers of recorded species for the biome Tropical and subtropical moist broadleaf forests, it is not possible to decide whether or not the Southeastern region actually has the highest species richness or whether the North region of Brazil has been underestimated due to limited access to new areas for collection of plant species.

Figure 2. Distribution of collection spots in Brazil (891313 occurrences of terrestrial plants covering 29712 species).

3.1.5. Relationships between species richness and range-restricted species richness

Endemism has already been combined with species richness to result in an endemic species richness indicator (Kier and Barthlott, 2001; Kier et al., 2009, Crisp et al., 2001). In this context the concepts of species richness and range-restricted species were combined to result in a biodiversity indicator called range-restricted species richness. The existence of a relationship between the total species richness and the range-restricted species richness was tested by using the fraction of range-restricted species (FRS) and the total species richness for all biomes and ecoregions. FRS is found to vary from 1 to 13% at ecoregion level (see Figure 3 and Table S1, ESM-1). For more than 50% of ecoregions – mainly the ones with small number of species – the FRS values range between 1 and 3%, and the results presented in Figure 3 suggest a weak correlation (correlation coefficients below 0.562 for Spearman rank-order and Pearson correlation tests, P < 0.05, see Table S3a, ESM-1) between the total number of species in a certain ecoregion and the number of range-restricted species.

The weak relationships between species richness and range-restricted species richness suggest that the use of a single biodiversity indicator must be considered with caution. Conserving an area of high biodiversity in terms of plant species richness does not guarantee that the largest number of

range-restricted species will be conserved. Furthermore, despite the correlation between plant species richness and area shown in Section 3.1.4, even small areas with a low number of species may hold a high number of range-restricted species, which is likely attributable to other factors such as the isolation of an area with special living conditions (e.g. islands) (Kruckeberg and Rabinowitz, 1985). The identification of range-restricted species from a more continuous mapping of species occurrences, using fine GIS grid cell resolutions instead of ecoregion or biome differentiations, would eliminate interference from the shape of a region and the presence of species at the boundaries. It would, however, also increase uncertainties due to the lack of data, especially in areas with comparatively few collection spots, like in the Amazon forest. Alternatively, the definition of range-restricted species used in this study could be adjusted to also include species that are present only in transition zones between ecoregions, which may thus not be captured in our definition. Using different species richness indicators in combination may contribute to an effective species conservation, with a combined focus on extinction of unique species (through the range-restricted-based indicator) and preservation of the overall diversity of occurring species (through the total species richness indicator).
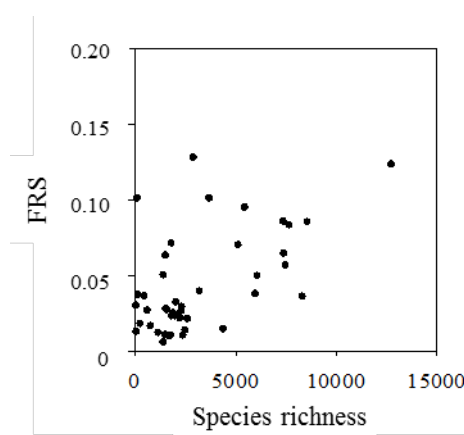

Figure 3. Relationships between the fraction of range-restricted species (FRS) and total species richness (correlation coefficients < 0.562 and P < 0.05). Statistical analysis data given in Table S3 in ESM-1.

3.2. Relationships between plant species richness and soil pH

3.2.1. Overall species richness and soil pH

Figure 4 and Figure S3 (ESM-1) show the distribution of terrestrial plant species richness as a function of soil pH at biome and whole country levels, respectively. At biome level, a very strong correlation was found with R^2 above 0.920 regardless of which type of regression fit (log normal or logistic) was used (see Figure 4). Lognormal distribution curves showed a slightly better fit to the

collected data (higher R^2) for biomes and were chosen as the preferred approach for this study. At ecoregion and whole-Brazil levels, soil pH and species richness also demonstrate statistically significant correlations, with R^2 ranging 0.830-1.000 for 40 out of 45 ecoregions (two additional ecoregions could not be described by the regression model) and $R^2 = 0.999$ for Brazil as a country, assuming a lognormal distribution model (see Table S3, ESM-1).

Data availability clearly influences the regression analysis, as demonstrated by the observation that the remaining five ecoregions with weaker correlations have relatively small data sets. A limited number of species occurrence data points thus results in uncertain boundaries for the pH ranges, potentially leading to species richness distribution curves with several peaks.

The pH range and the lognormal distributions observed for all spatial resolution levels are consistent with the general pattern of physiological tolerance behavior of plants observed for different environmental conditions (Pärtel, 2002). The lognormal distribution curves observed for nearly all regions indicate that for each ecoregion or biome, there is an optimum pH associated with the highest species richness within the region. Disregarding other environmental factors that may co-vary with soil pH, it indicates that decreasing or increasing pH from the optimum points may lead to a reduction in the number of species that occur. This interpretation should however be cautioned by the fact that the composition of the ecosystems may change as a result of change in soil pH; for example, a decreasing pH in an ecoregion with an initial pH above optimum would lead to an increase in total number of species, although the species composition and nature of the ecosystems may have dramatically changed. Such possible changes could not be included in the current study; they nevertheless constitute a topic worth exploring in future research. The distribution of the optimum pH across biomes and ecoregions in Brazil is further discussed in Section 3.3.
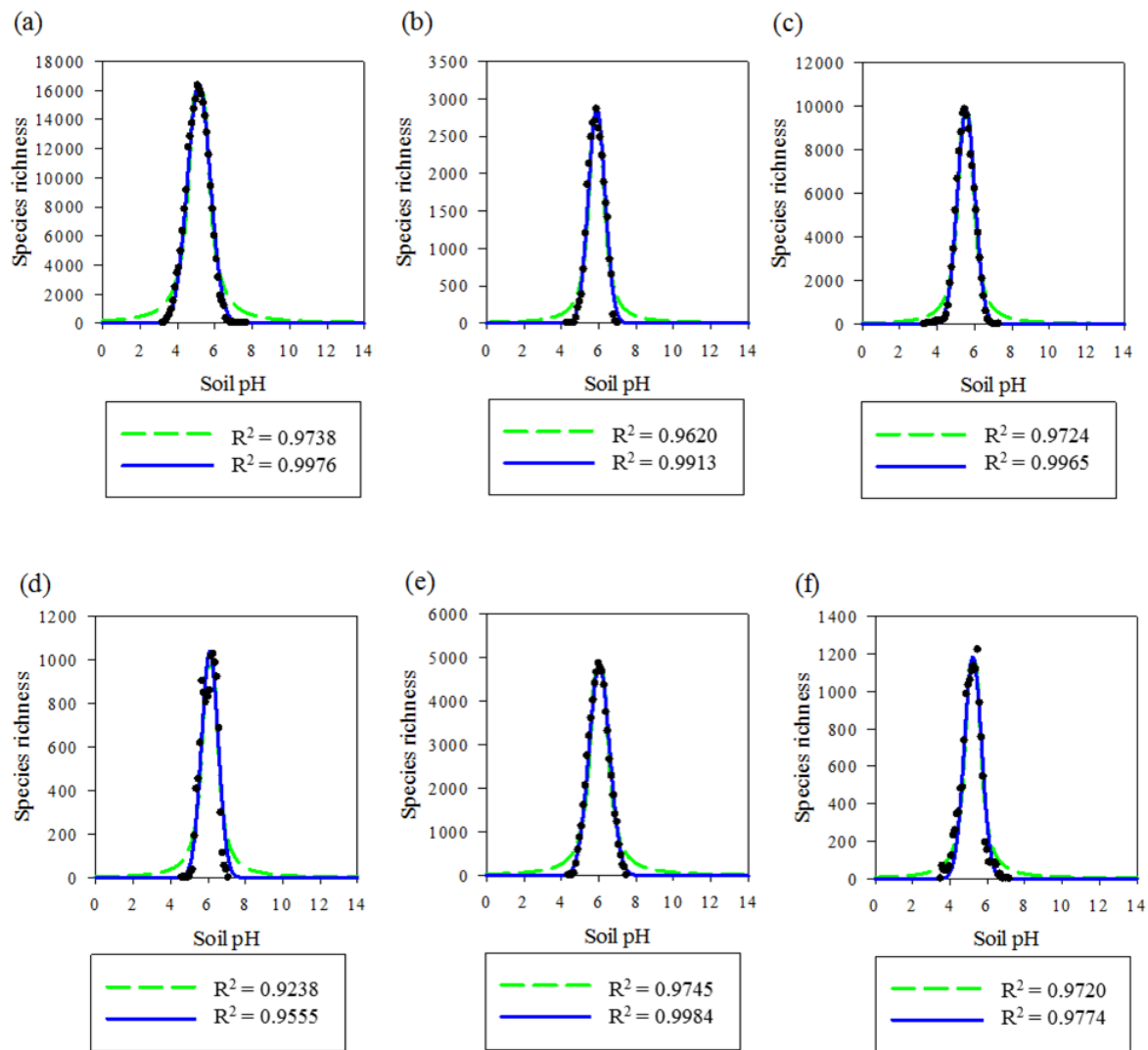
Figure 4. Comparison between continuous logistic regression (in green dotted curve) and lognormal regression (in blue thick curve) for the six Brazilian biomes: (a) Tropical and subtropical moist broadleaf forests, (b) Tropical and subtropical dry broadleaf forests, (c) Tropical and subtropical grasslands, savannas and shrublands, (d) Flooded grasslands and savannas, (e) Deserts and xeric shrublands and (f) Mangroves. Dots represent the collected data (number of species present at each 0.1 unit of pH).

3.2.2. Range-restricted species richness and soil pH

The correlations between the range-restricted species richness and the soil pH at country, biome and ecoregion levels are shown in Table S4 in ESM-1. The two variables demonstrate statistically significant correlations, with $R^2 = 0.982$ at country level (see Figure S3, ESM-1), R^2 ranging 0.855-0.995 at biome level (except for one biome, Figure 5), and R^2 ranging 0.700-0.995 for 32 out of 41 ecoregions (six ecoregions could not be described by the regression model; data not shown). When comparing with the total species richness distributions (Section 3.2.1), the range-restricted species richness distributions present slightly decreased R^2 values for most biomes, except for Mangroves

($R^2 = 0.583$). Out of the 41 ecoregions with regression results, only five ecoregions had R^2 lower than 0.500 due to small and unrepresentative data sets. This observation is consistent with previous results from Section 3.2.1, where ecoregions with lowest numbers of data points showed poor correlations.
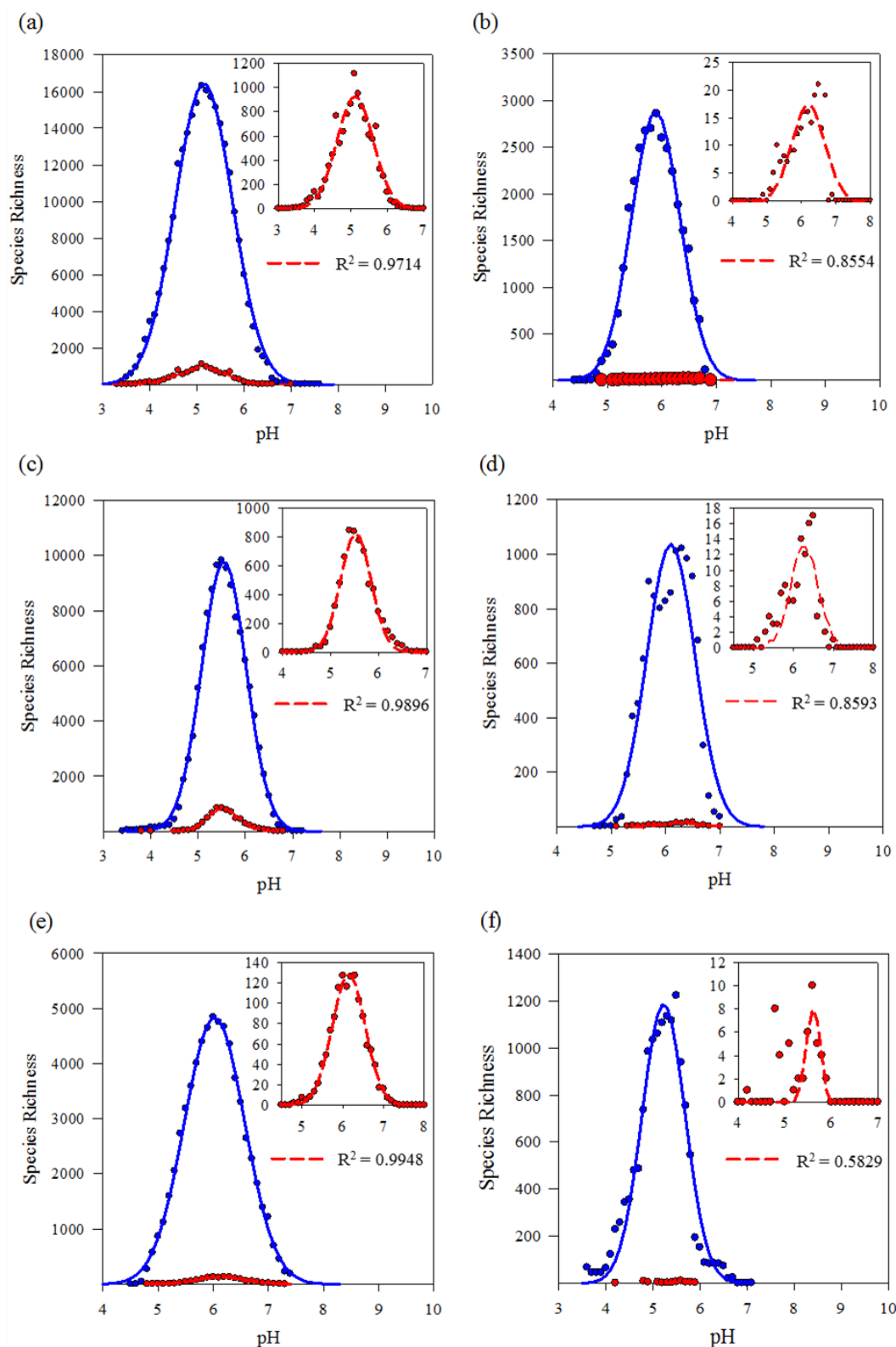

Figure 5. Species richness distribution for the entire list of species (thick curve, in blue) and for range-restricted species (dotted curve in red) for the six Brazilian biomes (inserts show distributions

for range-restricted species in higher resolution): (a) Tropical and subtropical moist broadleaf forests, (b) Tropical and subtropical dry broadleaf forests, (c) Tropical and subtropical grasslands, savannas and shrublands, (d) Flooded grasslands and savannas, (e) Deserts and xeric shrublands and (f) Mangroves. Dots represent the collected data (number of species present at each 0.1 unit of pH).

3.3. Variability of the species richness distributions across ecoregions

Within a biome, different ecoregions show different patterns in terms of optimum pH for the species richness distribution. Two distinct behaviors were observed among biomes in Brazil. Some biomes may include several ecoregions with different profiles, such as the Tropical and subtropical moist broadleaf forests, with 34 ecoregions and a high variability of optimum pH across the grouped ecoregions (see Figure 6a). Others may include few ecoregions and be well represented by just one or two of them (see Figure 6b, 6c and 6d). The disparities of species richness distributions and optimum pH values that can be observed at the ecoregion level are thus not observed at the biome level, reflecting an averaging effect across the larger areas and more variable conditions, as observed in Figures 6 and 7.

Pärtel and co-workers have previously demonstrated that, for low latitudes, the pool of plant species suited for low-pH soil (pH < 5.5) is larger than the pool of species suited for high-pH soil (pH > 5.5) (Pärtel, 2002; Pärtel et al., 2004). This is confirmed by our findings at country level, with entire Brazil having an optimum pH of 5.2. At the levels of biomes (Figure 7a) and ecoregions (Figure 7b), it can be observed that some specific biomes and ecoregions are associated with optimum pH above 5.5, hence showing a larger pool of species suited for high pH soil.

However, vegetation development and soil properties have a complex relationship that also depends on external stress factors, like droughts or fires (e.g. Folster et al. 2001). Our study furthers the understanding of these relationships and of the vegetation development at large, supporting biodiversity assessment. Yet, the influence of external factors like climate change and fires are not considered here, and are recommended to be investigated in future works.
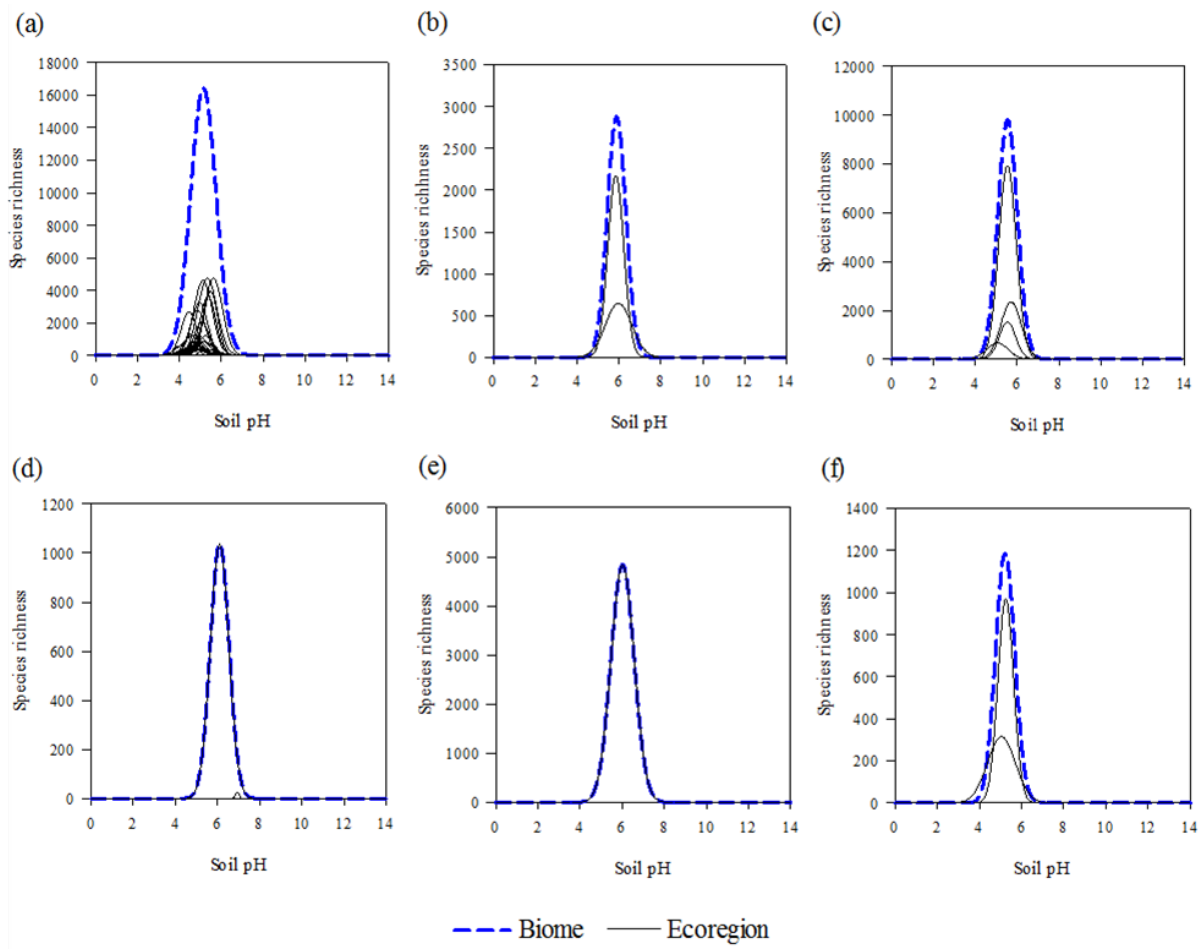
Figure 6. Species richness distribution: grouping of ecoregions within each biome (a) Tropical and subtropical moist broadleaf forests (incl. 34 ecoregions), (b) Tropical and subtropical dry broadleaf forests (incl. two ecoregions), (c) Tropical and subtropical grasslands, savannas and shrublands (incl. six ecoregions), (d) Flooded grasslands and savannas (incl. two ecoregions), (e) Deserts and xeric shrublands (incl. one ecoregion) and (f) Mangroves (incl. two ecoregions).
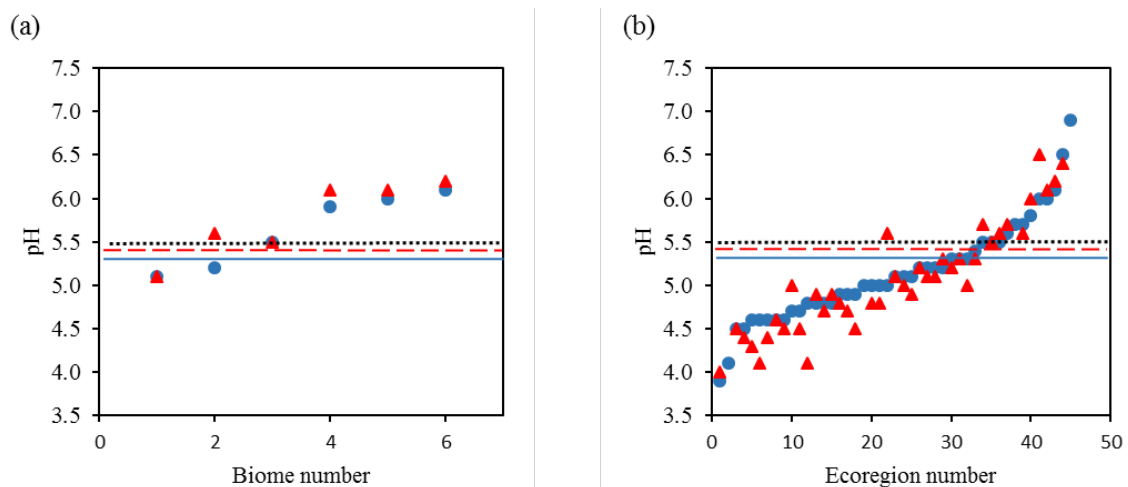

Figure 7. Optimum pH distribution at (a) biome level (total of 6 biomes) and (b) ecoregion level (total of 45 ecoregions). Circles (in blue) represent optimum pH for total species; Triangles (in red)

represent optimum pH for range-restricted species; Thick line (in blue) represents optimum pH for total species in Brazil; and dashed line (in red) represents optimum pH for range-restricted species in Brazil. Dotted line (in black) represents a boundary between acidic and non-acidic soils, with pH=5.5 (Pärtel, 2002; Pärtel et al., 2004).

4. CONCLUSIONS AND IMPLICATIONS FOR BIODIVERSITY ASSESSMENT

An inventory of terrestrial plants species is provided for Brazil, listing 29712 species with a harmonized nomenclature, spatially differentiated into biomes and ecoregions and identification of range-restricted species (species only occurring in one ecoregion). Range-restricted species richness is found to correlate poorly with total species richness and hence proposed as a complementary indicator of biodiversity. Besides preserving the diversity of species, the former also focuses on the preservation of unique species, avoiding their disappearance in the considered region. For future work it is suggested to compare the list of range-restricted species with the lists of endemics and The IUCN Red List of Threatened Species (IUCN, 2017) since endemism and threatened species can also be used as indicators of biodiversity that prioritize unique species to the region and indicate potential risks of extinction, respectively.

Additionally, the integration of physicochemical properties of the environment, in this case soil pH, contributes to the analysis of specificities within ecosystems that may be related to species vulnerability, supporting the identification of areas of high conservation priority. Regardless of the spatial resolution, the species richness variation as a function of soil pH indicates that decreasing or increasing pH from the optimum pH may be associated with a reduction in the number of species that occur. The relationship between terrestrial plant species richness and soil pH can thus support the assessment of impacts related to terrestrial acidification (Crespo-Mendes et al., 2018). The substantial empirical data presented here offer additional opportunities to estimate species loss per unit of land use that can be used to assess land use impacts. Such developments can help improve sustainability assessment approaches, such as Life Cycle Assessment, and support biodiversity conservation through planning and management of soil usage.

Mendes N.C., Laurent A., Bruun H.H., Hauschild M.Z., 2019. Relationships of terrestrial plant species richness with soil pH at the level of biome and ecoregion in Brazil. *Ecological Indicators* 98, 266-275.

List of references

Azevedo, L.B., van Zelm, R., Hendriks, A.J., Bobbink, R., Huijbregts, M.A., 2013. Global assessment of the effects of terrestrial acidification on plant species richness. Environmental pollution. 174, 10-15.

Boyle, B., Hopkins, N., Lu, Z., Garay, J.A.R., Mozzherin, D., Rees, T., Matasci, N., Narro, M.L., Piel, W.H., Mckay, S.J., Lowry, S., Freeland, C., Peet, R.K., Enquist, B.J., 2013. The taxonomic name resolution service: An online tool for automated standardization of plant names. BMC Bioinformatics. DOI: 10.1186/1471-2105-14-16.

Brazil Flora G, 2015. Brazilian Flora Checklist - Brazilian Flora 2020 project - Projeto Flora do Brasil 2020. Instituto de Pesquisas Jardim Botanico do Rio de Janeiro. Dataset/Checklist. doi:10.15468/1mtkaw. http://www.gbif.org/dataset/aacd816d-662c-49d2-ad1a-97e66e2a2908 (accessed 14 October 2015).

Brooks, T.M., Mittermeier, R.A., Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D., Rodrigues, A.S.L., 2006. Global Biodiversity Conservation Priorities. Science. 313, 58-61. DOI: 10.1126/science.1127609.

Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society (Series B). 345, 101-118.

Cooper, M., Mendes, L.M.S., Silva, W.L.C., Sparovek, G., 2005. A National Soil Profile Database for Brazil Available to International Scientists. Soil Science Society of America Journal. 69, 649-652. doi:10.2136/sssaj2004.0140.

Costa, D.P., Peralta, D.F., 2015. Bryophytes diversity in Brazil. Rodriguésia. 66, 1063-1071.

Crawley, M.J., Johnston, A.E., Silvertown, J., Dodd, M., Mazancourt, C., Heard, M.S., Henman, D.F., Edwards, G.R., 2005. Determinants of Species Richness in the Park Grass Experiment. The American Naturalist. 165, 179–192.

Crespo-Mendes, N., Laurent, A., Hauschild, M.Z., 2018. Effect factors of terrestrial acidification in Brazil for use in Life Cycle Impact Assessment. The International Journal of Life Cycle Assessment (accepted for publication).

Crisp, M.D., Laffan, S., Linder, H.P., Monro, A., 2001. Endemism in the Australian flora. Journal of Biogeography. 28, 183–198. doi:10.1046/j.1365-2699.2001.00524.x.

Mendes N.C., Laurent A., Bruun H.H., Hauschild M.Z., 2019. Relationships of terrestrial plant species richness with soil pH at the level of biome and ecoregion in Brazil. *Ecological Indicators* 98, 266-275.

Dinerstein, E., Olson, D.M., Graham, D.J., Webster, A.L., Primm, S.A., Bookbinder, M.P., Ledec, G., 1995. A Conservation Assessment of the Terrestrial Ecoregions of Latin America and the Caribbean. The World Bank, Washington.

Duprè, C., Stevens, C.J., Ranke, T., Bleeker, A., Peppler-Lisbach, C., Gowing, D.J.G., Dise, N.B., Dorland, E., Bobbink, R., Diekmann, M., 2010. Changes in species richness and composition in European acidic grasslands over the past 70 years: The contribution of cumulative atmospheric nitrogen deposition. Global Change Biology. 16, 344-357. https://doi.org/10.1111/j.1365-2486.2009.01982.x.

Ellis, E., Ramankutty, N., 2008. Putting people in the map: anthropogenic biomes of the world. Frontiers in Ecology and the Environment. 6, 439-447. doi:10.1890/070062.

Flann, C, editor. Global Compositae Checklist, 2009. www.compositae.org/checklist (accessed 14 October 2015).

Folster, H., Dezzeo, N., Priess, J.A., 2001. Soil-vegetation relationship in base-deficient premontane moist forest-savanna mosaics of the Venezuelan Guayana. Geoderma. 104, 95-113.

Forzza, R.C., Baumgratz, J.F.A., Bicudo, C.E.M., Canhos, D.A.L., Carvalho, A.A., Coelho, M.A.N., Costa, A.F., Costa, D.P., Hopkins, M.G., Leitman, P.M., Lohmann, L.G., Lughadha, E.N., Maia, L.C., Martinelli, G., Menezes, M., Morim, M.P., Peixoto, A.L., Pirani, J.R., Prado, J., Queiroz, L.P., Souza, S., Souza, V.C., Stehmann, J.R., Sylvestre, L.S., Walter, B.M.T., Zappi, D.C., 2012. New Brazilian Floristic List Highlights Conservation Challenges. BioScience. 62, 39-45. https://doi.org/10.1525/bio.2012.62.1.8.

GBIF, 2015. GBIF Annual Report 2014, Copenhagen: Global Biodiversity Information Facility, 34 pp. http://www.gbif.org/resource/annual_report_2014.

GBIF Secretariat: GBIF Backbone Taxonomy. doi:10.15468/39omei. http://www.gbif.org/species/6 (accessed 2 October 2015).

GBIF.org, 2015a - GBIF Occurrence Download. http://doi.org/10.15468/dl.jdhpbp (accessed 2 October 2015).

GBIF.org, 2015b - GBIF Occurrence Download http://doi.org/10.15468/dl.zo7qrt (accessed 2 October 2015).

GBIF.org, 2015c - GBIF Occurrence Download http://doi.org/10.15468/dl.zp3pxh (accessed 2 October 2015).

Mendes N.C., Laurent A., Bruun H.H., Hauschild M.Z., 2019. Relationships of terrestrial plant species richness with soil pH at the level of biome and ecoregion in Brazil. *Ecological Indicators* 98, 266-275.

GBIF.org, 2015d - GBIF Occurrence Download. http://doi.org/10.15468/dl.uo4mp1 (accessed 2 October 2015).

GBIF.org, 2015e - GBIF Occurrence Download. http://doi.org/10.15468/dl.0zlyfc (accessed 2 October 2015).

Gentry, A.H., 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. Annals of the Missouri Botanical Garden. 75, 1-34.

Gough, L., Shaver, G.R., Carroll, J., Royer, D.L., Laundre, J.A., 2000. Vascular plant species richness in Alaskan arctic tundra: The importance of soil pH. Journal of Ecology. 88, 54-66. https://doi.org/10.1046/j.1365-2745.2000.00426.x.

Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling. 157, 89-100. https://doi.org/10.1016/S0304-3800(02)00204-1.

Hengl, T., de Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Gonzalez, M.R., 2014. SoilGrids1km — Global Soil Information Based on Automated Mapping. PLoS ONE. 9, e105992. doi:10.1371/journal.pone.0105992.

Hoekstra, J.M., Boucher, T.M., Ricketts, T.H., Roberts, C., 2005. Confronting a biome crisis: global disparities of habitat loss and protection. Ecology Letters. 8, 23-29. doi: 10.1111/j.1461-0248.2004.00686.x.

IBGE - Instituto Brasileiro de Geografia e estatística, 2015. http://mapas.ibge.gov.br/ (accessed 09 September 2015).

IBGE/MMA, 2004. Mapa de Biomas do Brasil - Primeira Aproximação. http://www.ibge.gov.br

ILDIS - International Legume Database and Information Service, 2015. http://www.ildis.org/ (accessed 14 October 2015).

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

Kier, G., Kreft, H., Lee, T.M., Jetz, W., Ibisch, P.L., Nowicki, C., Mutkea, J., Barthlott, W., 2009. A global assessment of endemism and species richness across island and mainland regions. Proceedings of the National Academy of Sciences of the United States of America. 106, 9322-9327. https://doi.org/10.1073/pnas.

Mendes N.C., Laurent A., Bruun H.H., Hauschild M.Z., 2019. Relationships of terrestrial plant species richness with soil pH at the level of biome and ecoregion in Brazil. *Ecological Indicators* 98, 266-275.

Kruckeberg, A.R., Rabinowitz, D., 1985. Biological Aspects of Endemism in Higher Plants. Annual Review of Ecology and Systematics. 16, 447-479.

Longino, J.T., Coddington, J., Colwell, R.K., 2002. The Ant Fauna of a Tropical Rain Forest: Estimating Species Richness Three Different Ways. Ecology. 83(3), 689-702. https://doi.org/10.2307/3071874.

Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J., Chilquillo, E., Rønsted, N., Antonelli, A., 2015. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? Global Ecology and Biogeography. 24, 973-984. doi:10.1111/geb.12326.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I., White, E.P., 2007. Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. Ecology Letters. 10, 995-1015. https://doi.org/10.1111/j.1461-0248.2007.01094.x.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature. 403, 853-858. https://doi.org/10.1038/35002501.

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

Olson, D.M., Dinerstein, E., 1998. The Global 200: A representation approach to conserving the Earth's most biologically valuable ecoregions. Conservation Biology. 12, 502-515.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J., Owens, I.P.F., 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature. 436, 1016-1019. https://doi.org/10.1038/nature03850.

Prado, J., Sylvestre, L.D.S., Labiak, P.H., Windisch, P.G., Salino, A., Barros, I.C.L., Hirai,R.Y., Almeida, T.E., Santiago, A.C.P., Kieling-Rubio, M.A., Pereira, A.F.N., Øllgaard, B., Ramos, C.G.V., Mickel, J.T., Dittrich, V.A.O., Mynssen, C.M., Schwartsburd, P.B., Condack, J.P.S., Pereira, J.B.S., Matos, F.B., 2015. Diversity of ferns and lycophytes in Brazil. Rodriguesia. 66, 1073-1083. https://doi.org/10.1590/2175-7860201566410.

Mendes N.C., Laurent A., Bruun H.H., Hauschild M.Z., 2019. Relationships of terrestrial plant species richness with soil pH at the level of biome and ecoregion in Brazil. *Ecological Indicators* 98, 266-275.

Pärtel, M., Helm, A., Ingerpuu, N., Reier, U., Tuvi, E.L., 2004. Conservation of Northern European plant diversity: the correspondence with soil pH. Biological Conservation. 120, 525-531.

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

Roem, W.J., Berendse, F., 2000. Soil acidity and nutrient supply ratio as possible factors determining changes in plant species diversity in grassland and heathland communities. Biological Conservation. 92, 151-161. https://doi.org/10.1016/S0006-3207(99)00049-X.

Stevens, C.J., Duprè, C., Dorland, E., Gaudnik, C., Gowing, D.J.G., Bleeker, A., Diekmann, M., Alard,D., Bobbink, R., Fowler, D., Corcket, E., Mountford, J.O., Vandvik, V., Aarrestad, P.A., Muller, S., Dise, N.B., 2010. Nitrogen deposition threatens species richness of grasslands across Europe. Environmental Pollution. 158, 2940-2945. https://doi.org/10.1016/j.envpol.2010.06.006.

IUCN, 2017. The IUCN Red List of Threatened Species. Version 2017-1. http://www.iucnredlist.org

The Plant List, 2015. Version 1.1. Published on the Internet. http://www.theplantlist.org/ (accessed 14 October 2015).

Tropicos.org, 2015. Missouri Botanical Garden, St. Louis, MO, USA. http://www.tropicos.org (accessed 14 October 2015).

Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A., 2003. Neutral theory and relative species abundance in ecology. Nature. 424, 1035-1037. https://doi.org/10.1038/nature01883.

Zappi, D.C., Filardi, F.L.R., Leitman, P., Souza, V.C., Walter, B.M.T., Pirani, J.R., Morim, M.P., Queiroz, L.P., Cavalcanti, T.B., Mansano, V.F., Forzza, R.C., 2015. Growing knowledge: An overview of Seed Plant diversity in Brazil. Rodriguesia. 66, 1085-1113. https://doi.org/10.1590/2175-7860201566411.