Inescapable Taxonomists: Workable Biodiversity Management Must Base on a Minimum Field Work

Ariane Mirabel[1][*]
Eric Marcon[1]
Bruno Hérault[2]

Abstract
Résumé de l'article.

Keywords
mot-clés, séparés par des virgules

[1]UMR EcoFoG, AgroParistech, CNRS, Cirad, INRA, Université des Antilles, Université de Guyane.
    Campus Agronomique, 97310 Kourou, France.
[2]INPHB (Institut National Ploytechnique Félix Houphoüet Boigny)
    Yamoussoukro, Ivory Coast
*Corresponding author: ariane.mirabel@ecofog.gf, http://www.ecofog.gf/spip.php?article47

Contents

1. Introduction

The variety of tree species, their assemblages in space and their dynamics in time are determinant for forests ecological functions and productivity (**??**). It is crucial to maintain this diversity, especially in tropical forests where trees diversity is as threatened as it is valuable and unexplored (**?**). Maintain the diversity of tropical ecosystems requires setting protection areas and implementing sustainable forest management correctly calibrated through the assessment of spatial and temporal biodiversity patterns and their determinants (**?????**).

In that respect a large panel of diversity indices have been developed, constituting a refine framework catching all aspects of communities' diversity (**??**). Here we used the family of q-generalized or Tsallis entropy, converted into equivalent number of species (**???**). The diversities of this family derive from a unique formula, modulated by an order q that is the power to which is raised species frequency:

$$^qD = \sum_{i=1}^{N} \left(p_i^q \right)^{\frac{1}{1-q}}$$

In this formula q is the order of the diversity and p_i the relative abundance of species i in a community of N species. The order q determines the weight of species abundance in the metric: the higher the order, the higher the emphasis on common vs. rare species. A range of order q browses different balance between richness and evenness: for $q = 0$ the formula retrieves species richness, for $q = 1$ this is Shannon diversity that equally accounts for the richness and evenness components and for $q = 2$ this is the Simpson diversity as can be understood as the diversity of common species.

These tools though are based on forest inventories that require significant time, costs, and logistic (**?**). To respond these difficulties it has been proposed to reduce the inventory effort in focusing on some DBH, height classes or some particular taxa, or by leading inventories at family or genus level. These methods efficiently reflected biodiversity patterns at regional scales and along wide ecological gradients (**????**) but they did not clearly disentangled the richness and evenness components of diversity. They would miss the needs of studies at smaller time and spatial scale which involve faint variation, particularly in the Neotropics where the diversity of tree species is particularly large and complex (**?, ; ?; ?**). These restricting methods besides miss the needs of functional and phylogenetic approaches that are workable only with botanical names, to comply with phylogenetic trees and global trait database. Other methods proposed were rather interested in optimizing the sampling method of inventoried areas (**??**) but even then only small areas (under 1ha) were reasonably practicable because of the cost of high taxonomic resolution and the estimation of diversity remained uncertain, biased and the effects on richness and equitability still entangled.

It was also proposed to use inventories in vernacular names because they easier to attribute, more widely known and do not require vouchers collection or posterior botanical identification. Inventories in vernacular names besides already encompass the numerous forest inventories that logging companies conduct to quantify the available resources (????). The use of vernacular names successfully highlighted biodiversity patterns at regional scale and along large ecological gradients (?) but a clear and precise framework still need be developed. Vernacular names bear significant taxonomic uncertainty because they may correspond to several taxon, which besides depends of the field team (?). This taxonomic uncertainty should be correctly accounted for in the measure of diversity, as done by ?. They developed a framework based on Monte-Carlo process using the association between botanical and vernacular names, estimated through both general taxa-abundance table (?) and reference field botanical inventories. Their framework successfully rendered the ranking of plots diversity, their results came from the study of a large environmental gradient and highly different communities (??). We took up this framework to refine it and adapt it to the study of smaller time and spatial scales. The new framework proposed besides accounts for the specificities of the field teams and the studied community. It can also be applied whenever ratio of botanical determination in the initial inventory, be it partly or fully in vernacular names. In some cases indeed only the commercial or most recognizable species are directly identified at species level. This handling of several determination ratio could also give interesting application like promoting basic botanical formation for field workers to ensure a minimum determination ratio: this could be a small investment that highly increases the value of logging inventories for ecological surveys. Specifically we *(i)* revised the propagation Bayesian framework to account for both general taxa-association tables and reference field inventories. We extended the framework to analyze partly identified forest inventories and explored the impact of the different ratio of botanical determination on the performance of diversity estimators. *(ii)* We calibrated the framework using six 6.25ha plots from a tropical rainforest stand in French Guyana and determined the ideal balance between general taxa-association table and reference field inventories. We then determined the minimal sampling effort to be invested in the reference field inventories to have a correct estimation of diversity.

2. Methods

2.1 Study community and floristic flora
We based our analyses on a reference field inventory of a community of six 6.25ha rainforest plots in French Guyana. The inventory comes from the Paracou Research Station in French Guiana (5°18'N and 52°53'W) which is located in a lowland tropical rain forest with a dominance of Fabaceae, Chrysobalanaceae, Lecythidaceae and Sapotaceae. Mean annual precipitation averages $2980 mm.y^{-}1$ (30-y period) with a 3-months dry season ($< 100 mm.months-1$) from mid-August to mid-November and a one-month dry season in March (?). Elevation ranges between 5 and 50 m and mean annual temperature is 26°C. Soils correspond to thin acrisols over a layer of transformed saprolite with low permeability generating lateral drainage during heavy rains (?). We used the inventory of year 2015 of the six permanent plots of undisturbed forest (6.25ha each) settled within a 400-ha area. Trees with a DBH above 10cm are located and identified first with a vernacular name assigned by the field team before the following identification campaign when a botanist assigns a botanical species. The community represents 22904 trees belonging to 375 species and 63 families, identified by 290 different vernacular names. The initial taxonomic uncertainty that is the proportion of trees not fully identified concerned 3% of the community.

2.2 Propagation Framework
Our framework estimates the diversity of an inventory with a Monte-Carlo scheme that generates fully determined communities, by means of the association between vernacular names and the N botanical names of the inventory. This association is modelled by a multinomial distribution on the s_N botanical species. For any vernacular name of the inventory the distribution of associated botanical names followed a multinomial distribution $M([s_1, s_2, \ldots, s_N], [\alpha_1, \alpha_2, \ldots, \alpha_N])$, where $[s_1, s_2, \ldots, s_N]$ were the species recorded in the inventory and $[\alpha_1, \alpha_2, \ldots, \alpha_N]$ the corresponding probability of association. This process generates fully determined inventories when applied to all the initial vernacular names and repeated 1000 times it returns the estimation of the diversity and its uncertainty.

The probabilities $[\alpha_v]$ were determined with a Bayesian framework that could include two different dataset. First we could account for prior information from experts' knowledge in the form of a taxa-association table listing all botanical species likely corresponding to each vernacular name. The information of this table was summed up by a vector $[\lambda_v]$ where $\lambda_i = {}^1/m_v$ for the mv species which links to v was established and $\lambda_i = {}^\varepsilon/_{N-n_{table}}$ for others. The parameter ε stands for a background noise, it was set to 0.01. Second we could account for reference field inventory that give the observed association frequency between botanical and vernacular names. This inventory gives the vector $[\phi_v]$, where ϕ_i is the frequency to which a tree with the vernacular name v belonged to the botanical name i. We also added here a background noise, $\lambda_i = {}^\varepsilon/_{N-n_{field}}$ for the trees where no association was observed with nfield the number of associated botanical names and ε a background noise set to 0.01. The two information λ^v and ϕ^v were combined in a Multinomial-Dirichlet scheme (?) to return the final $[\alpha_v]$ distribution. To test the two different dataset we balanced them with a weighting parameter in the formula. Assuming a distribution of $[\phi_i]^v$ conditionally to $[\alpha_i]^v$ we had $[\alpha_i^v]$:
$$\left[\alpha_i^y|_{(1-w)\lambda_i^y, w.\phi_i^y}\right] = Dirichlet\left((1-w)\phi_i^y + w.\lambda_i^y\right).$$
When w was null, only the reference field inventory was considered, when it was 0.5 both dataset weighted equally and when w was 1 only the general taxa-association table was considered.

2.3 Simulation of the uncertainty gradient and determination of ptimal framework and reference protocol

To determine the impact of the determination ratio

$$\left(\frac{\text{number of vernacular name}}{\text{number of trees}}\right),$$

we simulated gradient of determination ratio by removing the botanical identification of an increasing proportion of species in inventory of reference. In the initial inventory a Kendall test ($\tau = -0.46, p < 10^{-}16$) showed that the probability of a species to be undetermined in an inventory is negatively linked to its abundance. Therefore, to simulate the indetermination gradient we sampled the species according to their abundance: the probability p_i of species i to be "undetermined" in a simulation was $f_i^{-0.1}$, with f_i its frequency. We applied our framework along the gradient of indetermination ratio to calibrate the framework and specifically find the best balance w between general taxa-association tables and reference field inventories. We also determined the minimum sampling effort for the reference field inventory, in terms of number of trees required to infer a correct vector of association frequencies $[\phi_v]$. We tested a range of sampling effort from 500 to 22 000 trees randomly selected from the whole inventory to calculate $[\phi_v]$. All our simulations of a gradient of indetermination ratio and sampling effort were repeated 1 000 times. From these iteration we assessed the performance of the diversity estimators through the average estimation bias, measured as the difference between the estimation and the diversity of the reference inventory (**?**), and through the relative estimation error, measured as the 95% confidence interval. We restricted our analysis on the Richness, Shannon and Simpson diversities which informs about both community's richness and equitability. To validate the convergence of the model we first simulated 100 values and realized a bootstrap through independent randomized subsamples of 2 to 100 simulations .

3. Results

3.1 Impacts of undetermination ratios and ideal settings

When considering both general taxa-correspondence table and reference field inventory diversity estimator showed a positively bias that increased with the indetermination ratio (Figure 1(a)). The bias of the estimator was besides increasing with the order of diversity q. For the order $q = 0$ the estimation did not significantly differ from the initial value of the real inventory while for the order $q = 1$ the overestimation reached 45% of the initial diversity and for the order $q = 2$ the overestimation reached 57%. When only the general taxa-correspondence table is considered (Figure 1(b)) the richness was highly underestimated, it reached 50% when the whole inventory was in vernacular names, while the Shannon and Simpson indices were both significantly overestimated, their respective bias reaching 67% and 125%. When only the reference field inventory is considered (Figure 1(c)) there were still estimation biases but they did not exceed 15% for any order of diversity.

We performed a bootstrap of the 100 simulations that showed a stabilization of variances after 60 simulations. We therefore set the number of simulations at 60 in the final script (Figure **??**).

> r Fig2, out.width = '60%', echo=FALSE,fig.cap="Degradation along a taxonomic uncertainty gradient of diversity estimated from a reference field inventories of 2 000 trees. 95% envelopes of the Richness, Shannon and Simpson diversities calculated along an uncertainty gradient from 0 to 100% of undetermined species."

3.2 Calibrating the sampling effort

We simulated a gradient of sampling effort for the reference field inventory, in terms of number of trees required to infer a the vector of association frequencies $[\phi_v]$, to identify the minimum effort for reliable diversity estimators. We tested a range of sampling effort from 500 to 22 000 trees randomly selected from the whole inventory (Figure 2). The biases of the estimators decreased with the order of diversity. The richness of reference is not faithfully retrieved before recovering the whole inventory (Figure 2(a)) but the 95% confidence interval of the estimator does not exceed 7% of the initial diversity. The Shannon and Simpson estimators showed lower biases, the bias of Shannon estimator falls to 15% for 2 000 reference trees (Figure 2(b)) and the bias of Simpson estimator falls to 6% (Figure 2(c)).

4. Discussion

In this paper we present a method of inventory protocols to correctly propagate taxonomic uncertainty of vernacular name in the measure of forest diversity. Our method is based on a Bayesian process which base is the probability of association between vernacular and botanical names. The comparison of several methods to build the basic vernacular-botanical association vectors demonstrated that the biases and the variability of the diversity estimator were much lower when reference field inventories are used rather than general taxa-association table. From this conclusion, we determined the minimum number of trees for the reference field inventory. To this end we run the estimation method with a set of association vectors computed from an increasing number of trees: we found that reference inventories should be based on a minimum of 2 000 trees, which ensures no more than 10% error for Shannon diversity and 1% error for Simpson diversity. We did not obtain an unbiased estimator of species Richness but demonstrated a linear correlation between estimated richness.

4.1 Prior field inventories for reliable diversity estimations

We set up a framework providing reliable and accurate diversity estimations in handling the taxonomic uncertainty of vernacular names due to their multiple correspondences to botanical species. In the line of **?** our method propagates the taxonomic uncertainty to diversity estimators through a Bayesian framework that was either based on general taxa-correspondence tables or reference field inventories. We
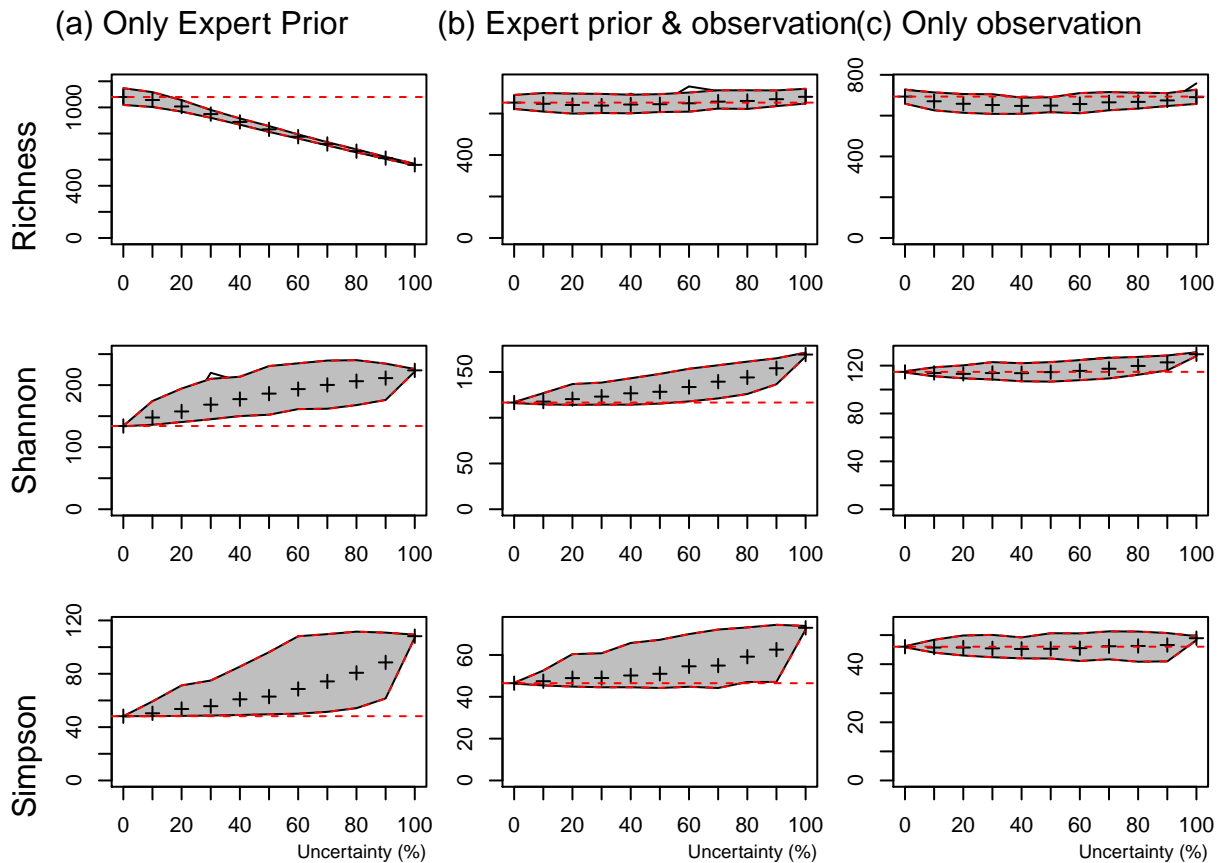
Figure 1. Indices degradation along a taxonomic uncertainty gradient. 95% envelopes of the Richness, Shannon and Simpson indices calculated through our propagation method along an uncertainty gradient from 0 to 100% of undetermined species. In (a) Only expert prior is considered to compute the association frequencies, in (b) both expert and observation prior are equally accounted for in the propagation method and in (c) only the observation prior is considered.
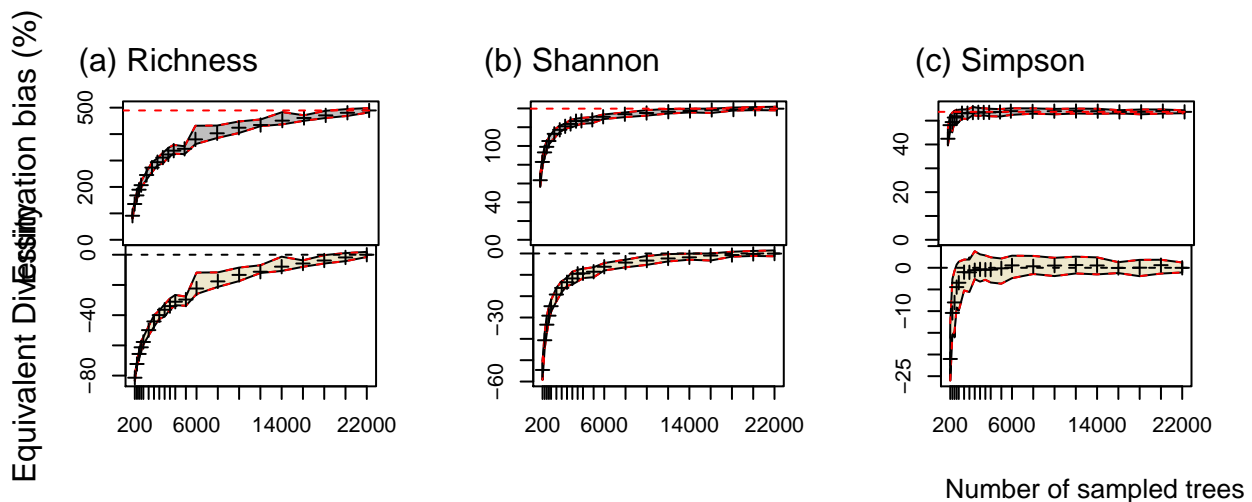

Figure 2. Degradation along a sampling effort gradient of the Richness, Shannon and Simpson diversities estimated for the reference inventory in vernacular names. The propagation method to estimate the diversities is only based on the reference field inventory. Above plots correspond to the estimated diversity in equivalent number of species and below plots correspond to the relative bias of the estimation compared to the value of the reference field inventory. For both dashed lines represent the value of the reference field inventory and crosses and red lines respectively represent the mean, 0.05 and 0.95 quantiles estimated after 1000 iterations.

determined the best balance between both dataset in applying our framework along a gradient of indetermination ratio ($\frac{\text{number of vernacular name}}{\text{number of trees}}$). Whenever its weight the account of the general taxa-association table underestimated the Richness and overestimated the Shannon and Simpson diversities (Figure 1). The use of general taxa-association table increases the equitability of the community in inflating the abundance of rare species at the expense of abundant ones. The association probabilities computed from this dataset are independent of species abundance so the vernacular name of a rare species will indifferently be associated to are or abundant species. The use of a reference field inventory to the contrary gives vector of association probability dependent of species abundance and retrieves the real abundance distribution so the diversity estimator however is much less biased. A reliable diversity estimator should then be based on reference field inventory performed beforehand by the working team in the studied community.

4.2 Calibration of the reference inventory

Reliable diversity estimator should be based on a reference field inventory ensuring small estimation biases and uncertainty. The determine the minimum size of this reference inventory we applied our framework for an gradient of sampling effort, in terms of number of trees used to compute the association probabilities. We found it difficult to retrieve the initial richness, as already suggested in previous analysis comparing random-sampling methods to these based on restricted inventories (**?**). However, if the estimator of richness was biased it varied little and should preserve the ranking of communities with similar indetermination ratio. This was coherent with various results linking the whole community richness to that of communities subsamples (**?**). The bias of the richness estimator though is a minor annoyance because for small time and spatial scale the richness is not necessary the most relevant diversity to consider (**????**). The estimator of Shannon and Simpson diversities were much less biased and reasonably. Our results showed that from 2 000 trees inventoried beforehand the Shannon and Simpson estimators respectively had 12% and 1% uncertainty.

5. Conclusion

A keystone for biodiversity conservation is to understand the determinants of communities assembly, especially in tropical forest where stands are as complex and species-rich as they are valuable and uncharted (**????**). Despite the study of tropical forests structure and composition however still a need large and intensive sampling in space and time which is hampered by the important costs of inventories in tropical forests (**?**). It is then urgent to develop methods alleviating the cost of inventories while producing accurate and unbiased estimation of the diversity. In that respect using vernacular names is promising because they are easier to attribute, known by the main part of field workers and proved to bear valuable information. Their reliability at genus level proved high but variable across tropical regions: it was estimated around 60-70% in French Guyana (**??**)

and from 32% to 67% in Central African (**?**). Although vernacular names can be used to estimate communities' diversity they generate significant taxonomic uncertainty due to the multiple correspondences between botanical and vernacular names. Reliably measure forest diversity from vernacular names requires setting adapted protocols and a propagation method of the taxonomic uncertainty. In this paper we present a method to propagate the taxonomic uncertainty of vernacular names to the measure of tropical forest diversity. We calibrated the corresponding inventory protocols to ensure accurate estimation: at the cost of an initial reference botanical inventory of 2 000 trees Shannon and Simpson diversities are estimated with respectively 10% and 1% accuracy. Our methods is besides workable to estimate functional and phylogenetic diversities indifferently. The method is based on reference field inventory and integrates the specificity of the working team and local forest structure. It is workable in all contexts and we propose that it be largely applied to get insights into the issue of vernacular names handling.

References

Baltanas, A. (2009) On the Use of Some Methods for the Estimation of Species Richness On the use of some methods for the estimation of species richness. *Oikos*, **65**, 484–492.

Baraloto, C., Hérault, B., Paine, C.E.T., Massot, H., Blanc, L., Bonal, D., Molino, J.F.F., Nicolini, E.a. & Sabatier, D. (2012a) Contrasting taxonomic and functional responses of a tropical tree community to selective logging. *Journal of Applied Ecology*, **49**, 861–870.

Baraloto, C., Molto, Q., Rabaud, S., Hérault, B., Valencia, R., Blanc, L., Fine, P.V.A. & Thompson, J. (2012b) Rapid Simultaneous Estimation of Aboveground Biomass and Tree Diversity Across Neotropical Forests: A Comparison of Field Inventory Methods ´. *Biotropica*, **0**, 1–11.

Berry, N.J., Phillips, O.L., Ong, R.C. & Hamer, K.C. (2008) Impacts of selective logging on tree diversity across a rainforest landscape: The importance of spatial scale. *Landscape Ecology*, **23**, 915–929.

Cannon, C.H. (1998) Tree Species Diversity in Commercially Logged Bornean Rainforest. *Science*, **281**, 1366–1368.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., A.Wardle, D., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S. & Naeem, S. (2012) Biodiversity loss and its impact on humanity. *Nature*, **489**, 326–326.

FAO (2014) Situation des Forêts du monde. Technical report, Organisation des Nations Unies pour l'alimentation et l'agriculture, Rome.

Feeley, K.J. & Silman, M.R. (2011) The data void in modeling current and future distributions of tropical species. *Global Change Biology*, **17**, 626–630.

Feldpausch, T.R., McDonald, A.J., Passos, C.A.M., Lehmann, J. & Riha, S.J. (2006) Biomass, harvestable area, and forest structure estimated from commercial timber inventories and remotely sensed imagery in southern Amazonia. *Forest Ecology and Management*, **233**, 121–132.

Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.a., Barlow, J., Peres, C.A., Bradshaw, C.J.a., Laurance, W.F., Lovejoy, T.E. & Sodhi, N.S. (2011) Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature*, **478**, 378–381.

Guitet, S., Cornu, J.F., Brunaux, O., Betbeder, J., Carozza, J.M. & Richard-Hansen, C. (2013) Landform and landscape mapping, French Guiana (South America). *Journal of Maps*, **9**, 325–335.

Guitet, S., Sabatier, D., Brunaux, O., Hérault, B., Aubry-Kientz, M., Molino, J.F.F. & Baraloto, C. (2014) Estimating tropical tree diversity indices from forestry surveys: A method to integrate taxonomic uncertainty. *Forest Ecology and Management*, **328**, 270–281.

Hawes, J.E., Peres, C.A., Riley, L.B. & Hess, L.L. (2012) Landscape-scale variation in structure and biomass of Amazonian seasonally flooded and unflooded forests. *Forest Ecology and Management*, **281**, 163–176.

Higgins, M.a. & Ruokolainen, K. (2004) Rapid Tropical Forest Inventory: a Comparison of Techniques Based on Inventory Data from Western Amazonia. *Conservation Biology*, **18**, 799–811.

Hill, M.O. (1973) Diversity and Evenness : A Unifying Notation and Its Consequences. *Ecological Society of America*, **54**, 427–432.

IUSS Working Group WRB (2015) World Reference Base for Soil Resources 2014, update 2015 Internatioanl soil classification system for naming soils and creating legends for soil maps. Technical report, FAO, ROME.

Jost, L. (2006) Entropy and diversity. *Oikos*, **113**, 363–375.

Keylock, C.J. (2005) Simpson diversity and the Shannon-Wiener index as special cases of a generalized entropy. **1**, 203–208.

Koh, L.P. & Sodhi, N.S. (2010) Conserving Southeast Asia's imperiled biodiversity: Scientific, management, and policy challenges. *Biodiversity and Conservation*, **19**, 913–917.

Magurran, A.E. (1988) *Ecological diversity and its measurement*. Princeton, nj, prince edition.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

McCarthy, M.A. (2007) *Bayesian Methods for Ecology*. Cambridge University Press.

Molino, J., Sabatier, D., Prévost, M., Frame, D., Gonzalez, S., Bilot-Guérin, V., De, M., Pêche, D., Molino, J., Sabatier, D., Prévost, M., Frame, D., Gonzalez, S., Amap, U.M.R. & Guyane, H.D. (2009) Etablissement d ' une liste des espèces d 'arbres de la Guyane Française. Technical report, IRD, Cayenne.

Oldeman, R. (1968) Sur la valeur des noms vernaculaires des plantes en Guyane Française. *Bois For Trop*, **117**, 17–23.

Phillips, O., Vasquez Martínez, R., Nunez Vargas, P., Monteagudo, A., Chuspe Zans, M., Galiano Sanchez, W., Pena Cruz, A., Timana, M., Yli-Halla, M. & Rose, S. (2003) Efficient plot-based floristic assessment of tropical forests. *Journal of Tropical Ecology*, **19**, 629–645.

Plumptre, A. (1996) Changes following 60 years of selective timber harvesting in the Budongo Forests Reserve, Unganda. *Forest Ecology and Management*, **89**, 101–113.

Pos, E.T., Guevara Andino, J.E., Sabatier, D., Molino, J.F., Pitman, N., Mogollón, H., Neill, D., Cerón, C., Rivas, G., Di Fiore, A., Thomas, R., Tirado, M., Young, K.R., Wang, O., Sierra, R., García-Villacorta, R., Zagt, R., Palacios, W., Aulestia, M. & ter Steege, H. (2014) Are all species necessary to reveal ecologically important patterns? *Ecology and Evolution*, **4**, 4626–4636.

Prance, G.T. (1994) A Comparison of the Efficacy of Higher Taxa and Species Numbers in the Assessment of Biodiversity in the A comparison of the efficacy of higher taxa and species numbers in the assessment of biodiversity in the neotropics. *Biological Sciences Biodiversity: Measurement and Estimation*, **345**, 89–99.

Purvis, A. & Hector, A. (2000) Getting the measure of biodiversity. *Nature Insight Biodiverstiy*, **405**, vol 405 no. 6783.

Réjou-Méchain, M., Fayolle, A., Nasi, R., Gourlet-Fleury, S., Doucet, J.L., Gally, M., Hubert, D., Pasquier, A. & Billand, A. (2011) Detecting large-scale diversity patterns in tropical trees: Can we trust commercial forest inventories? *Forest Ecology and Management*, **261**, 187–194.

Réjou-Méchain, M., Gourlet-fleury, S., Couteron, P., Nasi, R. & Thompson, J.D. (2008) Regional variation in tropical forest tree species composition in the Central African Republic : an assessment based on inventories by forest companies. *Journal of Tropical Ecology*, **24**, 663–674.

Sist, P., Pacheco, P., Nasi, R. & Blaser, J. (2015) Management of natural tropical forests for the future. Technical report, IUFRO WFSE.

Steege, H.T., Sabatier, D., Castellanos, H., Andel, T.V., Duivenvoorden, J., Oliveira, A.A.D., Ek, R., Lilwah, R., Maas, P., Mori, S., Ter Steege, H., Sabatier, D., Castellanos, H., Van Andel, T., Duivenvoorden, J., Adalardo De Oliveira, A., Ek, R., Lilwah, R., Maas, P. & Mori, S.

(2000) An analysis of the floristic composition and diversity of Amazonian forest including those of the Guiana Shield. *Journal of Tropical Ecology*, **16**, 801–828.

ter Steege, H., Pitman, N.C.a., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prévost, M.F., Spichiger, R., Castellanos, H., von Hildebrand, P. & Vásquez, R. (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**, 0–2.

Vellend, M., Lilley, P.L. & Starzomski, B.M. (2008) Using subsets of species in biodiversity surveys. *Journal of Applied Ecology*, **45**, 161–169.

Wagner, F., Hérault, B., Stahl, C., Bonal, D. & Rossi, V. (2011) Modeling water availability for trees in tropical forests. *Agricultural and Forest Meteorology*, **151**, 1202–1213.

Whittaker, R.H. (1972) Evolution and Measurement of Species Diversity. *Taxon*, **21**, 213–251.