

# Inescapable Taxonomists: Workable Biodiversity Management Based on a Minimum Field Work

Ariane Mirabel<sup>1\*</sup>

Eric Marcon<sup>1</sup>

Bruno Hérault<sup>2</sup>

## Abstract

Assess the fate of Neotropical forests requires to accurately measures are the base of reliable forest monitoring, crucial to assess the fate of neotropical forests in the current changing climate. The costs of botanical inventories and the taxonomic complexity of Neotropical forests make predominant forest inventories in vernacular names, although these hold high botanical uncertainty. Several methods proposed to compensate botanical uncertainties but none allowed reliable neither functional nor fine-scale diversity approaches. Here we offer a polyvalent diversity estimator propagating botanical uncertainties and workable in numerous specific cases. From a large neotropical inventory, we calibrated the estimator and through simulations we determined an ideal inventory protocol optimizing the costs and the accuracy of forest inventories. Our study first highlighted the unavoidable use to real inventories, compared to general vernacular/botanical tables, and the inescapable recourse to taxonomists to ensure robust diversity survey. Then our simulations allowed estimated the minimum sampling size (XX trees) and percentage of species accurately identified (80% of species known) for inventories to allow diversity estimations with a 10% error. The diversity estimator effectively assessed diversity for a variety of pre-logging and experimental forest monitoring, acknowledging the recourse to taxonomists, and enabled to design optimized inventory protocols.

## Keywords

Biodiversity Measurement, Tree Community, Neotropical Forests, Botanical Uncertainty Propagation, Bayesian Estimator

<sup>1</sup>UMR EcoFoG, AgroParistech, CNRS, Cirad, INRA, Université des Antilles, Université de Guyane.  
Campus Agronomique, 97310 Kourou, France.

<sup>2</sup>INPHB (Institut National Polytechnique Félix Houphouët Boigny)  
Yamoussoukro, Ivory Coast

\*Corresponding author: ariane.mirabel@ecofog.gf, <https://github.com/ArianeMirabel>

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methods</b>	<b>2</b>
2.1	Study community . . . . .	2
2.2	Diversity measures . . . . .	2
2.3	Diversity estimator . . . . .	2
2.4	Simulation of the uncertainty gradient and determination of optimal framework and reference protocol . . . . .	3
<b>3</b>	<b>Results</b>	<b>3</b>
3.1	Impacts of undetermination ratios and ideal settings	3
3.2	Calibrating the sampling effort . . . . .	3
<b>4</b>	<b>Discussion</b>	<b>5</b>
4.1	Prior field inventories for reliable diversity estimations	5
4.2	Calibration of the reference inventory . . . . .	5
<b>5</b>	<b>Conclusion</b>	<b>5</b>

## 1. Introduction

The variety of tree species, their assemblages in space and their dynamics in time are determinant of forests productivity and functioning (Magurran, 1988; Cardinale *et al.*, 2012). Preserve tree diversity is crucial to maintain forests functioning and services, specifically in hyper-diverse tropical forests where the biodiversity is as threatened as it is valuable and unexplored (Koh & Sodhi, 2010; Barlow *et al.*,

2018). Handling the conservation and management of tree diversity requires setting sensitive protection areas and sustainable forest management calibrated according to their spatial and temporal patterns and determinants (Margules & Pressey, 2000; Purvis & Hector, 2000; Gibson *et al.*, 2011; FAO, 2014; Sist *et al.*, 2015).

Correctly measure, map and manage forests biodiversity require accurate and large forest monitoring. The precision of forest inventories, though, is often limited by their significant cost in terms of time, money, and logistic (Feeley & Silman, 2011). Sampling methods were optimized to minimize these costs and maximize inventory accuracy. Some approaches would restrict inventories to some DBH or height classes, to specific taxa, or would opt for inventories at family or genus level. These methods efficiently translated biodiversity patterns at regional scales and along wide ecological gradients (Steege *et al.*, 2000; Higgins & Ruokolainen, 2004; Réjou-Méchain *et al.*, 2011; Pos *et al.*, 2014). However, these methods were either limited to small areas (under 1ha), sometimes remained biased or holding significant uncertainty, and usually proved limited to detect subtle diversity aspects and to disentangle richness from equitability parameters (Phillips *et al.*, 2003; Baraloto *et al.*, 2012b; Guitet *et al.*, 2014, ; Vellend *et al.*, 2008; Prance, 1994). Another approach proposed to use inventories in vernacular names instead of botanical species. Vernacular names indeed are easier to attribute, more common and usually do

not require vouchers collection or posterior botanical identification. Similarly though, vernacular names entail significant botanical uncertainties as the association with botanical species is usually highly variable in space and time (Oldeman, 1968). Besides, rough vernacular inventories would not suit the needs of functional and phylogenetic approaches that require identification at the botanical species to comply with phylogenetic and functional database. However the approach through vernacular names deserves further attention. First, it gives the opportunity to analyze pre-logging inventories conducted in large areas by logging companies. Second, as exhaustive inventories, they allow some post-process based on vernacular/botanical names association and allow the building of reliable diversity estimators (ter Steege *et al.*, 2006; Feldpausch *et al.*, 2006; Réjou-Méchain *et al.*, 2008, 2011). Following this idea Guitet *et al.* (2014) proposed a framework propagating vernacular names taxonomic uncertainties in diversity measures. The propagation framework was based on Monte-Carlo processes estimating forest diversity from the vernacular-botanical name association. These association combined prior information from both general taxa-abundance correspondence table (Molino *et al.*, 2009) and reference field inventories. The framework successfully rendered the ranking of plots diversity, but remained restricted to large environmental gradient and for highly different communities (Guitet *et al.*, 2014, 2013). In this study we offer to refine this framework and adapt it to diversity estimation at smaller spatial scales. The following diversity estimator is based on the specific case of the studied community and the inventory protocol. The diversity estimator besides suits all inventories whatever the ratio of botanical determination, *i.e.* ratio of vernacular compared to botanical names. It besides suits experimental specific as well as pre-logging inventories where only the commercial or most recognizable species are identified at species level.

Such diversity estimator allows maximizing the accuracy of diversity measures while minimizing the sampling effort, *i.e.* the size of inventoried communities and the number of accurately identified species. In this perspective we thought to calibrate an ideal inventory protocol optimized in terms of sampling effort and determination degree. From a real inventory, with complete vernacular and botanical identifications, we simulated ranges of sampling efforts and identification degrees along which we examined the bias and variability of the diversity estimator.

In this study we (i) redesigned a diversity estimator based on a Bayesian framework accounting for both general taxa-association tables and specific field inventories, and (ii) applied the estimator to a real Neotropical forest inventory to determine the sampling effort and determination degree of an ideal inventory protocol.

## 2. Methods

### 2.1 Study community

We based our analyses on the inventory of a Neotropical rainforest, from the Paracou Research Station in French Guiana (5°18'N and 52°53'W). The experimental site stands in a lowland tropical rainforest with a flora dominated by *Fabaceae*, *Chrysobalanaceae*, *Lecythidaceae* and *Sapotaceae*

families. Mean mean annual temperature is 26°C. and the mean annual precipitations average 2980mm.y<sup>-1</sup> (30-y period) with a 3-months dry season (< 100mm.months<sup>-1</sup>) from mid-August to mid-November and a one-month dry season in March (Wagner *et al.*, 2011). Elevation ranges between 5 and 50 m and soils correspond to thin acrisols over a layer of transformed saprolite with low permeability, generating lateral drainage during heavy rains (IUSS Working Group WRB, 2015). We used the 2015 inventory of six permanent plots of undisturbed forest (6.25ha each, 37.5ha inventoried in total). During inventories trees are identified first with a vernacular name assigned by the forest worker team, and afterward with a scientific name assigned by botanists during regular botanical campaigns. The community inventoried encompasses 22 904 trees belonging to 375 species and 63 families, identified by 290 different vernacular names. The initial taxonomic uncertainty was 3% of the community, *i.e.* the proportion of trees not identified with a botanical name.

### 2.2 Diversity measures

Among the large panel of diversity indices we examined here the family of q-generalized (Tsallis) entropy, widely adopted to assess all aspects of taxonomic, functional and phylogenetic diversities. The Tsallis diversity indices derive from a general formula, modulated by an order q emphasizing species frequency (1).

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

In the diversity formula, species relative abundance  $p_i$  in a community of  $N$  species is raised at the power  $q$  that is the order of the diversity. The higher the order  $q$ , the higher the emphasis on common vs. rare species, so browsing a range of order  $q$  corresponds assess a gradient balance between richness and evenness. The formula retrieves species richness for  $q = 0$ , Shannon diversity for  $q = 1$  where richness and evenness are equally accounted for and Simpson diversity, that can be understood as the diversity of common species, for  $q = 2$ . The Tsallis diversity indices would eventually be converted into equivalent number of species in our framework. The conversion in equivalent number of species, through Hill transformation, allows understandable analysis and comparisons among communities (Hill, 1973; Keylock, 2005; Jost, 2006).

### 2.3 Diversity estimator

The estimation framework is based on the distribution of the diversity of theoretical, fully determined communities. For each incomplete inventory, 1 000 theoretical inventories are simulated through the replacement by a Monte-Carlo scheme of vernacular names by botanical ones.

Theoretical inventories are simulated from the association probability between vernacular names and the  $N$  botanical names inventoried. This association is modelled for each vernacular name by a multinomial distribution on the  $s_N$  botanical species  $M([s_1, s_2, \dots, s_N], [\alpha_1, \alpha_2, \dots, \alpha_N])$ , where  $[s_1, s_2, \dots, s_N]$  are the species recorded in the inventory and

$[\alpha_1, \alpha_2, \dots, \alpha_N]$  their association probability with the vernacular name.

The probabilities  $[\alpha_v]$  were determined with a Bayesian framework based on the combination of botanical expertise and real, complete inventories. First, the estimation of  $[\alpha_v]$  was based on prior information from experts' knowledge in the form of a taxa-association table listing all botanical names likely corresponding to the vernacular name  $v$ . For a vernacular name  $v$  a vector  $[\lambda_v]$  giving the association probability  $\lambda_i = 1/m_v$  for  $v$  to be associated with the botanical name  $i$ . When no association between  $v$  and  $i$  was established,  $\lambda_i = \varepsilon / (N - n_{table})$ , with  $\varepsilon$  standing for a background noise set to 0.01 here. Second, the estimation of  $[\alpha_v]$  was based on reference field inventory giving the observed vernacular/botanical association frequency  $\phi_i$  constituting the vector  $[\phi_v]$ . Similarly, a background noise  $\varepsilon = 0.01$  was attributed to botanical names with no observed association with  $v$ , giving the association probability  $\lambda_i = \varepsilon / (N - n_{field})$ . The two vectors  $\lambda^v$  and  $\phi^v$  were combined in a Multinomial-Dirichlet scheme (McCarthy, 2007) to model the final  $[\alpha_v]$  distribution.

The relevance of the two parameters was tested in weighting their importance in the formula. With  $w$  the weighting parameter, assuming a distribution of  $[\phi_i]^v$  conditionally to  $[\alpha_i]^v$  we had (2)

$$[\alpha_i^v] : [\alpha_i^v | (1-w)\lambda_i^v, w\phi_i^v] = \text{Dirichlet}((1-w)\phi_i^v + w\lambda_i^v) \quad (2)$$

When  $w = 0$  only the reference field inventory was considered, when  $w = 0.5$  both dataset weighted equally and when  $w = 1$  only the general taxa-association table was considered.

## 2.4 Simulation of the uncertainty gradient and determination of optimal 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 (Baltanas, 2009), 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

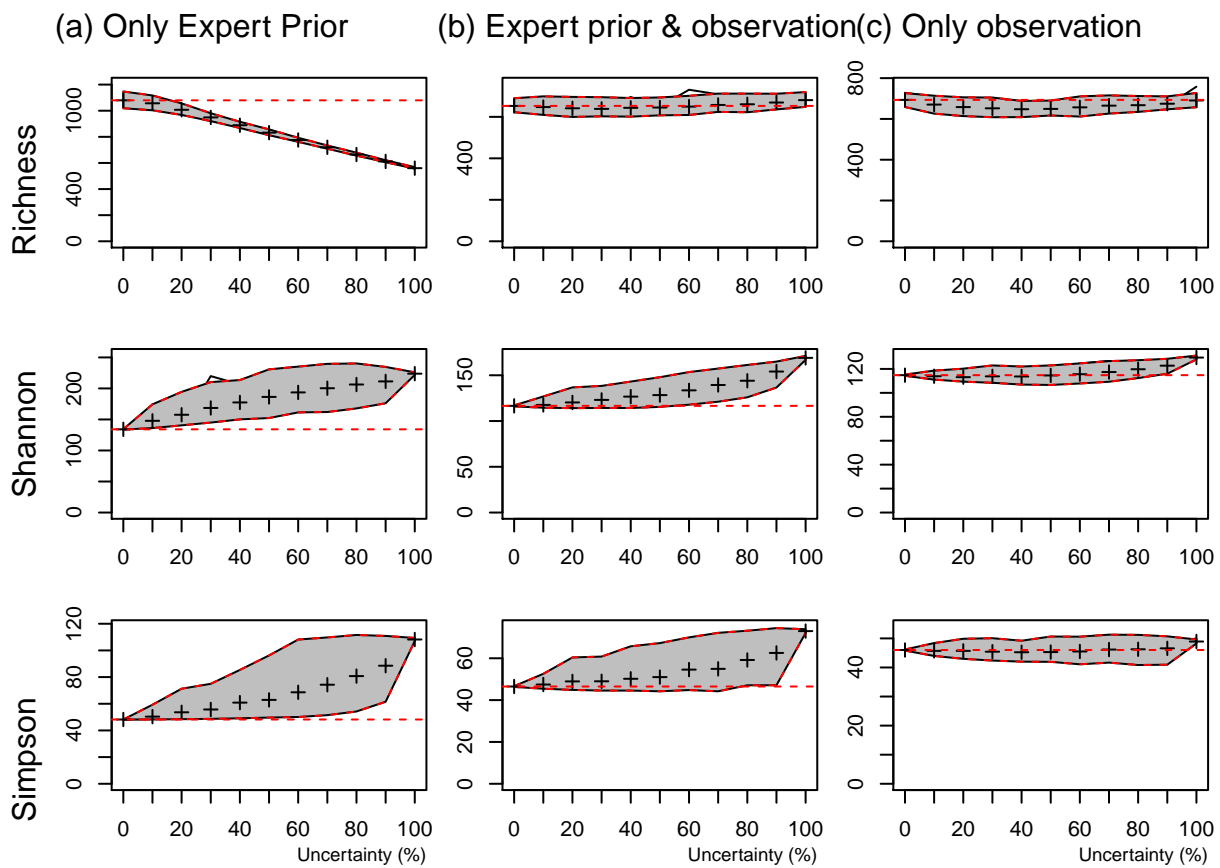
When considering both general taxa-correspondence table and reference field inventory diversity estimator showed a positively bias that increased with the indetermination ratio (Figure 1(a)). The bias of the estimator was besides increasing with the order of diversity  $q$ . For the order  $q = 0$  the 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



**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.



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 Guitet *et al.* (2014) 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 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.

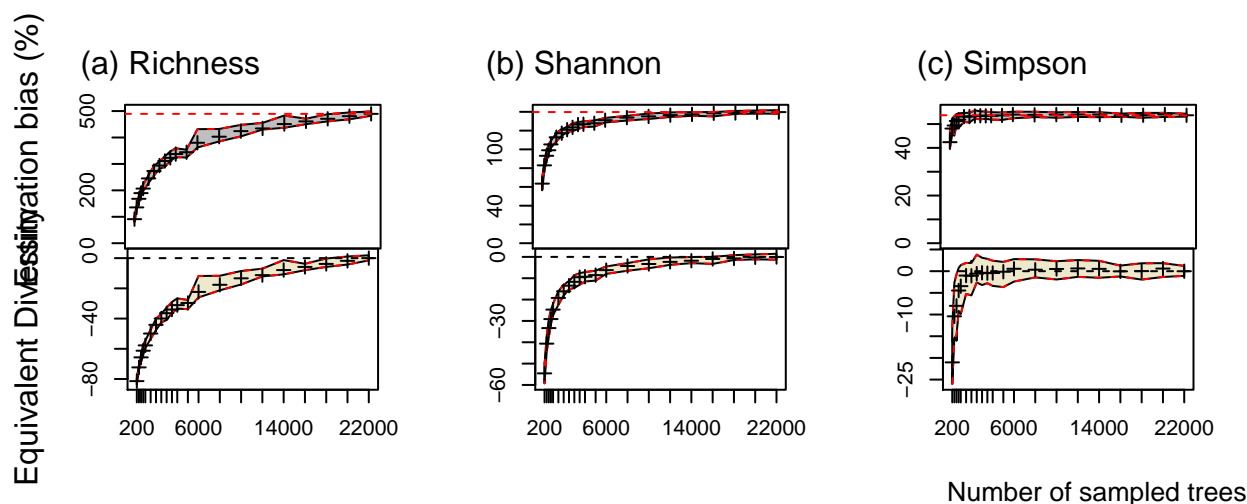
### 4.2 Calibration of the reference inventory

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 (Higgins & Ruokolainen, 2004). 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 (Vellend *et al.*, 2008). 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 (Baraloto *et al.*, 2012a; Berry *et al.*, 2008; Cannon, 1998; Plumptre, 1996). 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 (Magurran, 1988; Prance, 1994; Cardinale *et al.*, 2012; Sist *et al.*, 2015). 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 (Baraloto *et al.*, 2012b). 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 (Hawes *et al.*, 2012; Guitet *et al.*, 2014) and from 32% to 67% in Central African (Réjou-Méchain *et al.*, 2011). 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.



**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.

## 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.
- Barlow, J., França, F., Gardner, T., Hicks, C., Lennox, G., Berenguer, E., Castello, L., Economo, E., Ferreira, J., Guénard, B., Leal, C., Isaac, V., Lees, A., Paar, C., Wilson, S., Young, P. & Graham, N. (2018) The future of hyperdiverse tropical ecosystems. *Nature*.
- 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 International 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. *Oikos*, **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, Uganda. *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 Biodiversity*, **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.