

Generalization of the Partitioning of Shannon Diversity

Eric Marcon^{1*}, Ivan Scotti², Bruno Hérault³, Vivien Rossi³, Gabriel Lang^{4,5}

1 AgroParisTech, UMR Écologie des Forêts de Guyane, Kourou Cedex, France, **2** INRA, UMR Écologie des Forêts de Guyane, Kourou Cedex, France, **3** CIRAD, UMR Écologie des Forêts de Guyane, Kourou Cedex, France, **4** AgroParisTech, UMR 518 Math. Info. Appli., Paris, France, **5** INRA, UMR 518 Math. Info. Appli., Paris, France

Abstract

Traditional measures of diversity, namely the number of species as well as Simpson's and Shannon's indices, are particular cases of Tsallis entropy. Entropy decomposition, *i.e.* decomposing gamma entropy into alpha and beta components, has been previously derived in the literature. We propose a generalization of the additive decomposition of Shannon entropy applied to Tsallis entropy. We obtain a self-contained definition of beta entropy as the information gain brought by the knowledge of each community composition. We propose a correction of the estimation bias allowing to estimate alpha, beta and gamma entropy from the data and eventually convert them into true diversity. We advocate additive decomposition in complement of multiplicative partitioning to allow robust estimation of biodiversity.

Citation: Marcon E, Scotti I, Hérault B, Rossi V, Lang G (2014) Generalization of the Partitioning of Shannon Diversity. PLoS ONE 9(3): e90289. doi:10.1371/journal.pone.0090289

Editor: Jean Thioulouse, CNRS - Université Lyon 1, France

Received: January 25, 2013; **Accepted:** January 31, 2014; **Published:** March 6, 2014

Copyright: © 2014 Marcon et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work has benefited from an "Investissement d'Avenir" grant managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01). Funding came from the project Climfor (Fondation pour la Recherche sur la Biodiversité). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Co-author Bruno Hérault is a PLOS ONE Editorial Board member. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: eric.marcon@ecofog.gf

Introduction

Diversity partitioning means that, in a given area, the gamma diversity D_γ of all individuals found may be split into within (alpha diversity, D_α) and between (beta diversity, D_β) local assemblages. Alpha diversity reflects the diversity of individuals in local assemblages whereas beta diversity reflects the diversity of the local assemblages. The latter, D_β , is commonly derived from D_α and D_γ estimates [1]. Recently, a prolific literature has emerged on the problem of diversity partitioning, because it addresses the issue of quantifying biodiversity at large scale. Jost's push [2–5] has helped to clarify the concepts behind diversity partitioning but mutually exclusive viewpoints have been supported, in particular in a forum organized by Ellison [6] in *Ecology*. A recent synthesis by Chao *et al.* [7] wraps up the debate and attempts to reach a consensus. Traditional measures of diversity, namely the number of species as well as Simpson's and Shannon's indices, are all special cases of the Tsallis entropy [8,9]. The additive decomposition [10] of these diversity measures does not provide independent components but Jost [3] derived a non-additive partitioning of entropy which does.

A rigorous vocabulary is necessary to avoid confusion. *Unrelated* or *independent* (*sensu* [7]) means that the range of values of $^qD_\beta$ is not constrained by the value of $^qD_\alpha$, which is a desirable property. *Unrelated* is more pertinent than *independent* since diversity is not a random variable here, but *independent* is widely used, by [3] for example. We will write *independent* throughout the paper for convenience. We will write *partitioning* only when independent components are obtained and *decomposition* in other cases.

Tsallis entropy can be easily transformed into Hill numbers [11]. Jost [3] called Hill numbers *true diversity* because they are homogeneous to a number of species and have a variety of desirable properties that will be recalled below. We will call *diversity*

true diversity only, and *entropy* Simpson and Shannon indices as well as Tsallis entropy. The multiplicative partitioning of true γ diversity allows obtaining independent values of α and β diversity when local assemblages are equally weighted.

However, we believe that the additive decomposition of entropy still has something to tell us. In this paper, we bring out an appropriate mathematical framework that allows us to write Tsallis entropy decomposition. We show its mathematical equivalence to the multiplicative partitioning of diversity. This is simply a generalization of the special case of Shannon diversity [12]. Doing so, we establish a self-contained (*i.e.* it does not rely on the definitions of α and γ entropies) definition of β entropy, showing it is a generalized Jensen-Shannon divergence, *i.e.* the average generalized Kullback-Leibler divergence [13] between local assemblages and their average distribution. Beyond clarifying and making explicit some concepts, we acknowledge that this decomposition framework largely benefits from a consistent literature in statistical physics. In particular, we rely on it to propose bias corrections that can be applied to Tsallis entropy in general. After bias correction, conversion of entropy into true diversity provides independent, easy-to-interpret components of diversity. Our findings complete the well-established non-additive (also called pseudo-additive) partitioning of Tsallis entropy. We detail their differences all along the paper.

Methods

Consider a meta-community partitioned into several local communities (let $i = 1, 2, \dots, I$ denote them). n_i individuals are sampled in community i . Let $s = 1, 2, \dots, S$ denote the species that compose the meta-community, n_{si} the number of individuals of species s sampled in the local community i , $n_s = \sum_i n_{si}$ the total number of individuals of species s , $n = \sum_s \sum_i n_{si}$ the total number

of sampled individuals. Within each community i , the probability p_{si} for an individual to belong to species s is estimated by $\hat{p}_{si} = n_{si}/n_i$. The same probability for the meta-community is p_s . Communities may have a weight, w_i , satisfying $p_s = \sum_i w_i p_{si}$. The commonly-used $w_i = n_i/n$ is a possible weight, but the weighting may be arbitrary (e.g. the sampled areas).

We now define precisely entropy. Given a probability distribution $\mathbf{p}_s = \{p_1; p_2; \dots; p_s; \dots; p_s\}$, we choose an information function $\mathcal{I}(p_s)$, which is a decreasing function of p_s having the property $\mathcal{I}(1)=0$: information is much lower when a frequent species is found. Entropy is defined as the average amount of information obtained when an individual is sampled [14]:

$$H = \sum_s p_s \mathcal{I}(p_s) \quad (1)$$

The best-known information function is $\mathcal{I}(p_s) = -\ln(p_s)$. This defines the entropy of Shannon [15]. $\mathcal{I}(p_s) = (1-p_s)/p_s$ yields the number of species minus 1 and $\mathcal{I}(p_s) = 1-p_s$, Simpson's [16] index. Relative entropy is defined when the information function quantifies how different an observed distribution \mathbf{p}_s is different from the expected distribution \mathbf{p}'_s . The Kullback-Leibler [17] divergence is the best-known relative entropy, equal to $\sum_s p_s \ln(p_s/p'_s)$. Shannon's beta entropy has been shown to be the weighted sum of the Kullback-Leibler divergence of local communities, where the expected probability distribution of species in each local community is that of the meta-community [12,18]:

$$^1H_\beta = \sum_i w_i \sum_s p_{si} \ln\left(\frac{p_{si}}{p_s}\right) \quad (2)$$

Let us define γ as the meta-community's diversity, α as local communities' diversities, and β as diversity between local communities. Tsallis γ entropy of order q is defined as:

$$^qH_\gamma = \frac{1 - \sum_s p_s^q}{q-1} \quad (3)$$

and the corresponding α entropy in the local community i is:

$$^qH_{\alpha_i} = \frac{1 - \sum_s p_{si}^q}{q-1} \quad (4)$$

The natural definition of the total α entropy is the weighted average of local community's entropies, following Routledge [19]:

$$^qH_\alpha = \sum_i w_i ^qH_{\alpha_i} \quad (5)$$

This is the key difference between our decomposition framework and the non-additive one. Jost [3] proposed another definition, $^qH_\alpha = \sum_i (w_i^q / \sum_i w_i^q) ^qH_{\alpha_i}$, i.e. the normalized q -expectation of the entropy of communities [20] rather than their weighted mean. It is actually a derived result, see the discussion below. Our results rely on Routledge's definition (see Appendix S1).

α and γ diversity values are given by Hill numbers qD , called "numbers equivalent" or "effective number of species", i.e. the number of equally-frequent species that would give the same level of diversity as the data [14]:

$$^qD_\gamma = \left(\sum_s p_s^q \right)^{\frac{1}{1-q}} \quad (6)$$

Routledge α diversity is:

$$^qD_\alpha = \left(\sum_i w_i \sum_s p_{si}^q \right)^{\frac{1}{1-q}} \quad (7)$$

Combining (3) and (6) yields:

$$^qD_\gamma = (1 - (q-1) ^qH_\gamma)^{\frac{1}{1-q}} \quad (8)$$

We also use the formalism of deformed logarithms, proposed by Tsallis [21] to simplify manipulations of entropy. The deformed logarithm of order q is defined as:

$$\ln_q x = \frac{x^{1-q} - 1}{1-q} \quad (9)$$

It converges to \ln when $q \rightarrow 1$.

The inverse function of $\ln_q x$ is the deformed exponential:

$$e_q^x = [1 + (1-q)x]^{\frac{1}{1-q}} \quad (10)$$

The basic properties of deformed logarithms are:

$$\ln_q(xy) = \ln_q x + \ln_q y - (q-1)(\ln_q x)(\ln_q y) \quad (11)$$

$$\ln_q \frac{1}{x} = -x^{q-1} \ln_q x \quad (12)$$

$$e_q^{x+y} = e_q^x e_q^{\frac{y}{1-(q-1)x}} \quad (13)$$

Tsallis entropy can be rewritten as:

$$^qH_\gamma = \frac{1 - \sum_s p_s^q}{q-1} = - \sum_s p_s^q \ln_q p_s \quad (14)$$

Diversity and Tsallis entropy are transformations of each other:

$$^qH_\gamma = \ln_q ^qD_\gamma \quad (15)$$

$${}^qD_\gamma = e_q^{qH_\gamma} \quad (16)$$

Decomposing diversity of order q

We start from the multiplicative partitioning of true diversity.

$${}^qD_\gamma = {}^qD_\alpha {}^qD_\beta \quad (17)$$

If community weights are equal, β diversity is independent of α diversity (it is whatever the weights if α diversity is weighted according to Jost, but this is not our choice). We will consider the unequal weight case later.

β diversity is the equivalent number of communities, *i.e.* the number of equally-weighted, non-overlapping communities that would have the same diversity as the observed ones.

We want to explore the properties of entropy decomposition. We calculate the deformed logarithm of equation (17):

$$\ln_q {}^qD_\gamma = \ln_q {}^qD_\alpha + \ln_q {}^qD_\beta - (q-1)(\ln_q {}^qD_\alpha)(\ln_q {}^qD_\beta) \quad (18)$$

$$\Leftrightarrow {}^qH_\gamma = {}^qH_\alpha + \ln_q {}^qD_\beta - (q-1)({}^qH_\alpha)(\ln_q {}^qD_\beta) \quad (19)$$

Equation (19) is Jost's partitioning framework (equation 8f in [3]). Jost retains $H_B = \ln_q {}^qD_\beta$ as the β component of entropy partitioning. It is independent of ${}^qH_\alpha$ (they are respective transformations of independent ${}^qD_\beta$ and ${}^qD_\alpha$), contrarily to the β component of the additive decomposition [10,22] defined as ${}^qH_\gamma - {}^qH_\alpha$.

After some algebra requiring Routledge's definition of α diversity detailed in Appendix S1, we obtain from equation (19):

$${}^qH_\gamma - {}^qH_\alpha = \frac{\sum_i w_i \sum_s p_{si}^q - \sum_s p_s^q}{q-1} \quad (20)$$

The right term of equation (20) is a possible definition of the β component of additive decomposition. It can be much improved if we consider $\sum_s p_s^q = \sum_s p_s^{q-1} \sum_i w_i p_{si}$ and rearrange equation (20) to obtain:

$${}^qH_\gamma - {}^qH_\alpha = \sum_i w_i \sum_s p_{si}^q \ln_q \frac{p_{si}}{p_s} \quad (21)$$

We obtained the β entropy of order q . It is the weighted average of the generalized Kullback-Leibler divergence of order q (previously derived by Borland *et al.* [13] in thermodynamics) between each community and the meta-community:

$${}^qH_\beta = \sum_i w_i {}^qH_{\beta i} \quad (22)$$

$${}^qH_\beta = \sum_s p_{si}^q \ln_q \frac{p_{si}}{p_s} \quad (23)$$

${}^qH_\beta$ converges to the Kullback-Leibler divergence when $q \rightarrow 1$.

The average Kullback-Leibler divergence between several distributions and their mean is called Jensen-Shannon divergence [23], so our β entropy ${}^qH_\beta$ can be called *generalized Jensen-Shannon divergence*. It is different from the non-logarithmic Jensen-Shannon divergence [24] which measures the difference between the equivalent of our α entropy and $-\sum_i w_i p_{si}^q \ln_q p_s^q$ (the latter is not Tsallis γ entropy).

Our results are summarized in Table 1, including transformation of entropy into diversity. The partition of entropy of order q is formally similar to that of Shannon entropy. It is in line with Patil and Taillie's [14] conclusions: ${}^qH_\beta$ is the information gain attributable to the knowledge that individuals belong to a particular community, beyond belonging to the meta-community.

Information content of generalized entropy

Both ${}^qH_\gamma$ and ${}^qH_\beta$ must be rearranged to reveal their information function and explicitly write them as entropies. Straightforward algebra yields:

$${}^qH_\gamma = - \sum_s p_s \frac{p_s^{q-1} - 1}{q-1} \quad (24)$$

$${}^qH_\beta = \sum_s p_{si} \frac{p_{si}^{q-1} - p_s^{q-1}}{q-1} \quad (25)$$

The information functions respectively tend to those of Shannon entropy when $q \rightarrow 1$.

Properties of generalized β entropy

${}^qH_\beta$ is not independent of ${}^qH_\alpha$. Only Jost's H_B is an independent β component of diversity indices. But ${}^qH_\beta$ takes place in a generalized decomposition of entropy. Its limit when $q \rightarrow 1$ is Shannon β entropy, and in this special case only ${}^qH_\beta$ is independent of ${}^qH_\alpha$.

${}^qH_\beta$ is interpretable and self-contained (*i.e.* it is not just a function of γ and α entropies): it is the information gain brought by the knowledge of each local community's species probabilities related to the meta-community's probabilities. It is an entropy, defined just as Shannon β entropy but with a generalized information function.

${}^qH_\beta$ is always positive (proof in [25]), so entropy decomposition is not limited to equally-weighted communities.

Bias correction

Estimation bias (we follow the terminology of Dauby and Hardy [26]) is a well-known issue. Real data are almost always samples of larger communities, so some species may have been missed. The induced bias on Simpson entropy is smaller than on Shannon entropy because the former assigns lower weights to rare species, *i.e.* the sampling bias is even more important when q decreases.

We denote ${}^q\hat{H}$ the naive estimators of entropy, obtained by applying the above formulas to estimators of probabilities (such as ${}^q\hat{H}_\beta = \sum_s \hat{p}_{si}^q \ln_q (\hat{p}_{si}/\hat{p}_s)$). Let ${}^q\tilde{H}$ denote the estimation-bias corrected estimators. Chao and Shen's [27] correction can be

Table 1. Values of entropy and diversity for generalized entropy of order q and Shannon entropy.

Diversity measure	Generalized entropy	Shannon
γ entropy	${}^q H_\gamma = -\sum_s p_s^q \ln_q p_s$	${}^1 H_\gamma = -\sum_s p_s \ln p_s$
β entropy	${}^q H_\beta = \sum_i w_i \sum_s p_{si}^q \ln_q \frac{p_{si}}{p_s}$	${}^1 H_\beta = \sum_i w_i \sum_s p_{si} \ln \frac{p_{si}}{p_s}$
True γ diversity (Hill number)	${}^q D_\gamma = e_q^{H_\gamma}$	${}^1 D_\gamma = e^{H_\gamma}$
True β diversity (numbers equivalent)	${}^q D_\beta = e_q^{\frac{{}^q H_\beta}{1-q-1/H_\beta}}$	${}^1 D_\beta = e^{H_\beta}$

The deformed logarithm formalism allows presenting all orders of entropy as a generalization of Shannon entropy. Generalized β entropy is a generalized Kullback-Leibler divergence, i.e. the information gain obtained by the knowledge of each community's composition beyond that of the meta-community. Robust estimation of the entropy of real communities requires estimation bias correction introduced in the text.

doi:10.1371/journal.pone.0090289.t001

applied to all of our estimators. It relies on the Horvitz-Thomson [28] estimator which corrects a sum of measurements for missing species by dividing each measurement by $1-(1-\hat{p}_{si})^n$, i.e. the probability for each species to be present in the sample. Next, the sample coverage of community i , denoted C_i , is the sum of probabilities the species of the sample represent in the whole community. It is easily estimated [29] from the number of singletons (species observed once) of the sample, denoted S_i^1 , and the sample size n_i :

$$\hat{C}_i = 1 - \frac{S_i^1}{n_i} \quad (26)$$

The sample coverage of the meta-community is estimated the same way: $\hat{C} = 1 - S^1/n$. An unbiased estimator of p_{si} is $\tilde{p}_{si} = \hat{C}_i \hat{p}_{si}$, and $\tilde{p}_s = \hat{C} \hat{p}_s$. Combining sample coverage, Horvitz-Thomson and equation (23) estimator yields:

$${}^q \tilde{H}_\gamma = - \sum_s \frac{(\hat{C} \hat{p}_s)^q \ln_q \hat{C} \hat{p}_s}{1 - (1 - \hat{C} \hat{p}_s)^n} \quad (27)$$

$${}^q \tilde{H}_\beta = \sum_s \frac{(\hat{C}_i \hat{p}_{si})^q \ln_q \frac{\hat{C}_i \hat{p}_{si}}{\hat{C} \hat{p}_s}}{1 - (1 - \hat{C}_i \hat{p}_{si})^n} \quad (28)$$

Another estimation bias has been widely studied by physicists. The latter generally consider that all species of a given community are known and their probabilities quantified. Their main issue is not at all missing species but the non-linearity of entropy measures (see [30] for a short review). Probabilities p_s are estimated by \hat{p}_s . For $q > 0$, estimating p_s^q by $(\hat{p}_s)^q$ is an important source of underestimation of entropy. Grassberger [31] derived an unbiased estimator \tilde{p}_s^q under the assumption that the number of observed individuals of a species along successive samplings follows a Poisson distribution, as in Fisher's model [32] although arguments are different. Grassberger shows that:

$$\tilde{p}_s^q \approx n_s^{-q} \left(\frac{\Gamma(n_s + 1)}{\Gamma(n_s - q + 1)} + \frac{(-1)^n \Gamma(1 + q) \sin \pi q}{\pi(n + 1)} \right) \quad (29)$$

where $\Gamma(\cdot)$ is the gamma function ($\Gamma(n) = (n-1)!$ if n is an integer). Practical computation of $\Gamma(n_s + 1)$ is not possible for large samples so the first term of the sum must be rewritten as:

$\Gamma(n_s + 1)/\Gamma(n_s - q + 1) = \Gamma(q)/\mathcal{B}(n_s - q + 1, q)$ where \mathcal{B} is the beta function. This estimator can be plugged into the formula of Tsallis γ entropy to obtain:

$${}^q \tilde{H}_\gamma = \frac{1 - \sum_s \tilde{p}_s^q}{q - 1} \quad (30)$$

Other estimations of p_s^q are readily detailed here. Holste *et al.* [33] derived the Bayes estimator of p_s^q (with a uniform prior distribution of probabilities not adapted to most biological systems) and, recently, Hou *et al.* [34] derived ${}^2 \tilde{H}_\gamma = n/(n-1)(1 - \sum_s \hat{p}_s^2)$, namely the bias correction proposed by Good [29] and Lande [10]. Bonachela *et al.* [30] proposed a balanced estimator for not too small probabilities p_s which do not follow a Poisson distribution. This may be applied to low-diversity communities. In summary, the estimation of p_s^q requires assumptions about the distribution of p_s and Grassberger's correction is recognized by all these authors as the best up-to-date for very diverse communities. Better corrections exist but are available for special values of q only, such as the recent Chao *et al.*'s estimator of Shannon entropy [35].

The correction for missing species by Chao and Shen and that for non-linearity by Grassberger ignore each other. Chao and Shen's bias correction is important when q is small and becomes negligible for $q = 2$ while Grassberger's correction increases with q , vanishing for $q = 0$. A rough but pragmatic estimation-bias correction is the maximum value of the two corrections. It cannot be applied when $q < 0$ (Grassberger's correction is limited to positive values of q) neither to β entropy (Chao and Shen's correction can but Grassberger's can't). An estimator of β entropy will be obtained as the difference between unbiased γ and α entropy.

We illustrate this method with a tropical forest dataset already investigated by [12]. Two 1-ha plots were fully inventoried in the Paracou field station in French Guiana. This results in 1124 individual trees (diameter at breast height over 10 cm) belonging to 229 species. Figure 1 shows diversity values calculated for q between 0 and 2, with and without correction. Chao and Shen's bias correction is inefficient for $q > 1.5$ and can even be worse than the naive estimator. In contrast, Grassberger's correction is very good for high values of q , but ignores the missed species and decreases when $q \rightarrow 0$. The maximum value offers an efficient correction. By nature, α and γ diversity values decrease with q (proof in [36]): around 300 species are estimated in the meta-community ($q = 0$, Figure 1), but the equivalent number of species is only 73 for $q = 2$.

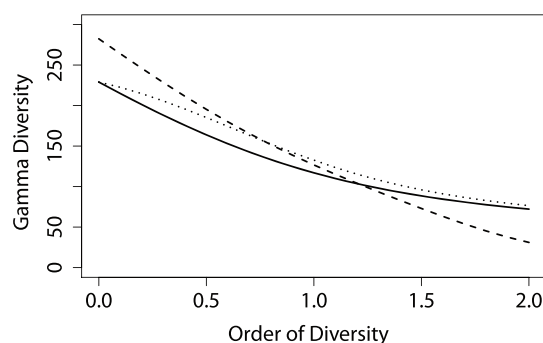


Figure 1. Profile of the γ diversity in a tropical forest meta-community. Data from French Guiana, Paracou research station, 2 ha inventoried, 1124 individual trees, and 229 observed species. Solid line: without estimation bias correction; dotted line: Grassberger correction; dashed line: Chao and Shen correction. The maximum value is our bias-corrected estimator of diversity.
doi:10.1371/journal.pone.0090289.g001

Converting unbiased entropy into diversity introduces a new bias issue because of the non-linear transformation by the deformed exponential of order q . We follow Grassberger's argument: this bias can be neglected because the transformed quantity (i.e. the entropy) is an average value (the information) over many independent terms, so it has little fluctuations (contrarily to the species probabilities whose non-linear transformation causes serious biases, as we have seen above).

We used Barro Colorado Island (BCI) tropical forest data [37] available in the vegan package [38] for R [39] to show the convergence of the estimators to the real value of diversity. 21457 trees were inventoried in a 50 hectare plot. They belong to 225 species. Only 9 species are observed a single time, so the sample coverage is over 99.99%. The inventory can be considered as almost exhaustive and used to test bias correction. We subsampled the BCI community by drawing chosen size samples (from 100 to 5000 trees) in a multinomial distribution respecting the global species frequencies. We drew 100 samples of each size, calculated their entropy, averaged it and transformed the result into diversity before plotting it in Figure 2. For low values of q , Chao and Shen's correction is the most efficient. It is close to the Chao1 estimator [40] of the number of species for $q=0$ (not shown). A correct estimation of diversity of order 0.5 is obtained with less than 1000 sampled trees (around 2 hectares of inventory). When q increases, Grassberger bias correction is more efficient: for $q=1.5$ and over, very small samples allow a very good evaluation. Both corrections are equivalent around $q=1.2$ (not shown).

Examples

Simple, theoretical example

We first propose a very simple example to visualize the decomposition of entropy. A meta-community containing 4 species is made of 3 communities C1, C2 and C3 with weights 0.5, 0.25 and 0.25. The number of individuals of each species in communities are respectively (25, 25, 40, 10), (70, 20, 10, 0), (70, 10, 0, 20). The resulting meta-community species frequencies is (0.475, 0.2, 0.225, 0.1). Note that community weights do not follow the number of individuals (100 in each community). No bias correction is necessary since the sample coverage is 1 in all cases. Entropy decomposition is plotted in Figure 3. For $q=0$, α and γ entropy equal the number of species minus 1. The meta-community's γ entropy is 3, including α entropy equal to 2.5

(the average number of species minus 1). β entropy is 0.5, equal to the averaged sum of communities contributions. C2's β entropy is negative (the total β entropy is always positive, but communities contributions can be negative).

Considering Shannon entropy, C1 is still the most diverse community (4 species versus 3 in C2 and C3, and a more equitable distribution: it has the greatest α entropy equal to 1.29). C2 and C3 have the same α entropy (their frequency distributions are identical) equal to 0.8. C3's species distribution is more different from the meta-community's than the others: it has the greatest β entropy equal to 0.34. Entropies can be transformed into diversities to be interpreted: the α diversity of communities is 3.6, 2.2 and 2.2 effective species, the total α diversity equals 2.8 effective species. The meta-community's γ diversity is 3.5 effective species (quite close to its maximum value 4 if all species were equally distributed) and β diversity is 1.2 effective communities: the same β diversity could be obtained with 1.2 theoretical, equally weighted communities with no species in common.

Real data application

We now want to compare diversity between Paracou and BCI, the two forests introduced in the previous section.

Diversity profiles are a powerful way to represent diversity of communities advocated recently by [36], as a function of the importance given to rare species which decreases with q . Comparing diversity among communities requires plotting their diversity profiles rather than comparing a single index since profiles may cross (examples from the literature are gathered in [36], Figure 2). Yet, estimation bias depends on the composition of communities, questioning the robustness of comparisons: a consistent bias correction over orders of entropy is required.

Entropy is converted to diversity and plotted against q in Figure 4 for our two forests: plots are given equal weight since they have the same size and gamma diversity is calculated for each meta-community. Paracou is more diverse, whatever the order of diversity. Bias correction allows comparing very unequally sampled forests (2 ha in Paracou versus 50 ha in BCI, sample coverage equal to 92% versus 99.99%).

β diversity profile is calculated between the two plots of Paracou. To compare it with BCI which contains 50 1-ha plots, we calculated α and β entropies between all couples of BCI plots, averaged them and converted them into β diversity (α and β entropies are required to calculate β diversity). We also calculated the 95% confidence envelope of β diversity between two 1-ha plots of BCI by eliminating the upper and lower 2.5% of the distribution of all plot couples β diversity. We chose to use Chao and Shen's correction up to $q=1.2$ and Grassberger's correction for greater q to obtain comparable results in the 1225 pairs of BCI plots. Figure 5 shows Paracou's β diversity is greater than BCI's, especially when rare species are given less importance: for $q=2$ (Simpson diversity), two plots in BCI are as different from each other as 1.2 plots with no species in common, while Paracou's equivalent number of plots is 1.7. In other words, dominant species are very different in Paracou plots, while they are quite similar on average between two BCI plots.

The shape of β diversity profiles is more complex than that of γ diversity. At $q=0$, β diversity equals the ratio between the total number of species and the average number of species in each community [7]. At $q=1$, it is the exponential of the average Kullback-Leibler divergence between communities and the meta-community. A minimum is reached between both. Over $q=1$, β diversity increases to asymptotically reach its maximum value equal to ${}^{\infty}D_{\gamma}$, i.e. the inverse of the probability of the most frequent species of the meta-community, divided by ${}^{\infty}D_{\alpha}$, i.e. the

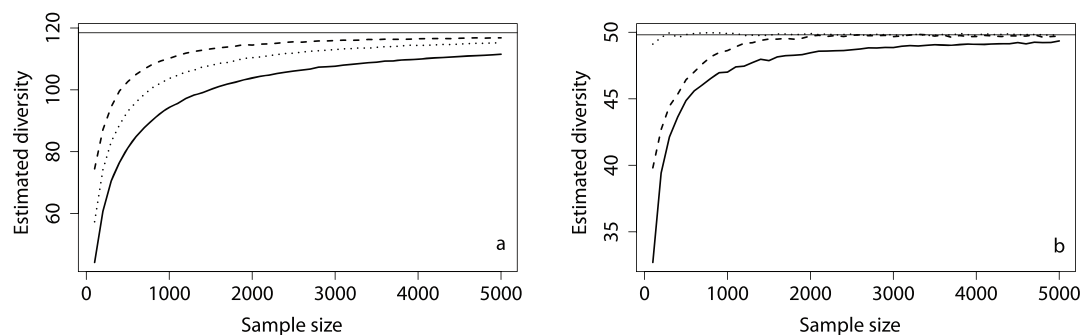


Figure 2. Efficiency of bias correction. Estimation of diversity of the BCI tropical forest plot for two values of the order of diversity q (a: 0.5, b: 1.5). The horizontal line is the actual value calculated from the whole data (around 25000 trees, species frequencies are close to a log-normal distribution). Estimated values are plotted against the sample size (100 to 5000 trees). Solid line: naive estimator with no correction; dotted line: Grassberger correction; dashed line: Chao and Shen's correction. For $q=0.5$, Chao and Shen perform best. For $q=1.5$, Grassberger's correction is very efficient even with very small samples.
doi:10.1371/journal.pone.0090289.g002

inverse of the probability of the most frequent species in each community.

Discussion

Diversity can be decomposed in several ways, multiplicatively, additively or non-additively if we focus on entropy. A well-known additive decomposition of Simpson entropy is as a variance (that of Nei [41] among others). It is derived in Appendix S2. It is not a particular case of our generalization: the total variance between communities actually equals β entropy but the relative contribution of each community is different. Among these several decompositions, only the multiplicative partitioning of equally-weighted communities (17) and the non-additive partitioning of entropy (19) allow independent α and β components (except for the special case of $q=1$), but unequal weights are often necessary and ecologists may not want to restrict their studies to Shannon diversity.

We clarify here the differences between non-additive partitioning and our additive decomposition and we address the question of unequally-weighted communities.

Additive versus non-additive decomposition

Jost [3] focused on independence of the β component of the partitioning. He showed (appendix 1 of [3]) that if communities are not equally weighted the only definition of ${}^qH_\alpha$ allowing independence between α and β components is ${}^qH_\alpha = \sum_i (w_i^q / \sum_i w_i^q) {}^qH_{\alpha_i}$. The drawback of this definition is that α may be greater than γ entropy if $q \neq 1$ and community weights are not equal. Each component of entropy partitioning can be transformed into diversity as a Hill number.

We have another point of view. We rely on Patil and Taillie's concept of diversity of a mixture (section 8.3 of [14]), which implies Routledge's definition of α entropy. It does not allow independence between α and β components of the decomposition except for the special case of Shannon entropy, but it ensures that

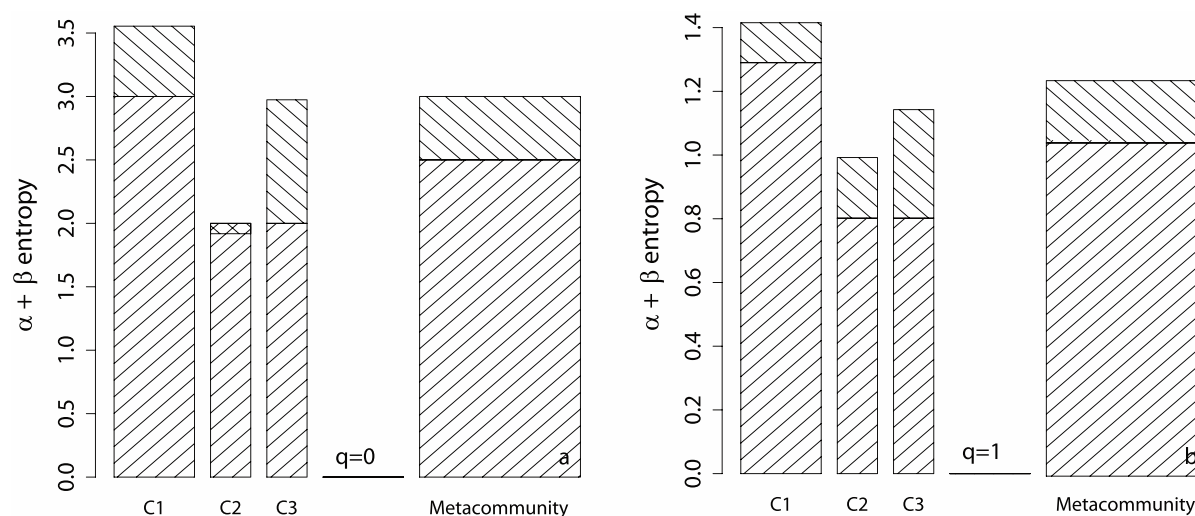


Figure 3. Decomposition of a meta-community entropy. The meta-community is made of three communities named C1, C2 and C3 (described in the text). Their α entropy ${}^qH_\alpha$ (bottom part of the bars) and their contribution to β entropy ${}^qH_\beta$ (top part of the bars) are plotted for $q=0$ (a) and $q=1$ (b). The width of bars is each community's weight. α and β entropies of the meta-community are the weighted sums of those of communities, so the area of the rectangles representing community entropies sum to the area of the meta-community's (width equal to 1). γ entropy of the meta-community is α plus β entropy.
doi:10.1371/journal.pone.0090289.g003

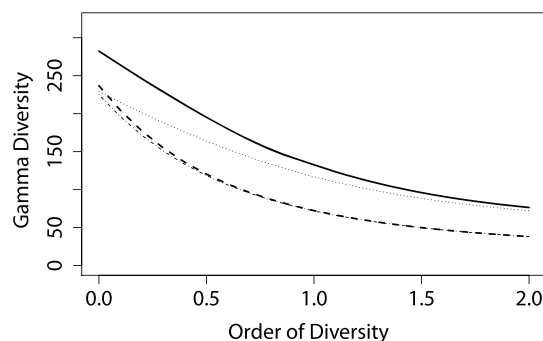


Figure 4. Paracou and BCI γ diversity. Diversity of the forest stations is compared. Solid line: Paracou with bias correction; dotted line: Paracou without bias correction; dashed line: BCI with bias correction; dotted dashed line: BCI without bias correction. Without bias correction, Paracou and BCI diversities appear to be similar for low values of q . Bias correction shows that Paracou is undersampled compared to BCI (actually around 1000 trees versus 25000). Paracou is much more diverse than BCI.
doi:10.1371/journal.pone.0090289.g004

β entropy is always positive. We believe that independence is not essential when dealing with entropy, as it emerges when converting entropy to diversity, at least when community weights are equal. The β component of the decomposition cannot be transformed into β diversity without the knowledge of α entropy but we have shown that it is an entropy, justifying the additive decomposition of Tsallis entropy.

The value of β entropy cannot be interpreted or compared between meta-communities as shown by [4], but combining α and β entropy allows calculating β diversity (Table 1).

Unequally weighted communities

Routledge's definition of α entropy does not allow independence between α and β diversity when community weights are not equal, and β diversity can exceed the number of communities [7]. We show here that the number of communities must be reconsidered to solve the second issue. We consider the independence question then.

We argue that Routledge's definition always allows to reduce the decomposition to the equal-weight case. Consider the example of Chao *et al.* [7]: two communities are weighted $w_1=0.05$ and $w_2=0.95$, their respective number of species are $S_1=100$ and $S_2=10$, no species are shared, and we focus on $q=0$ for simplicity. ${}^0D_\gamma$ equal 110 species, ${}^0D_\alpha$ is the weighted average of S_1 and S_2 equal to 14.5, so ${}^0D_\beta$ is 7.6 effective communities, which is more than the actual 2 communities. But this example is equivalent to that of a meta-community made of 1 community identical to the first one and 19 communities identical to the second one, all equally weighted. β diversity of this 20-community meta-community is 7.6 effective communities.

A more general presentation is as follows. A community of weight w can be replaced by any set of n identical communities of weights w_1, \dots, w_n provided that the sum of these weights is w , without changing α , β and γ diversity of the meta-community because of the linearity of Routledge's definition of entropy. Any unequally weighted set of community can thus be transformed into an equally weighted one by a simple transformation (strictly speaking, if weights are rational numbers).

Consider a meta-community made of several communities with no species in common, and say the smallest one (its weight is w_{min}) is the richest (its number of species is S_{max}). If S_{max} is large enough,

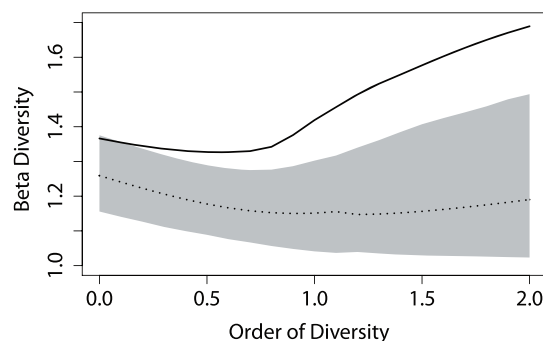


Figure 5. Paracou and BCI β diversity. β diversity profile between Paracou plots (solid line) is compared to that of any two plots of BCI (dotted line with 95% confidence envelope).
doi:10.1371/journal.pone.0090289.g005

the number of species of the meta-community is not much more than it (poor communities can be neglected). γ richness ${}^0D_\gamma$ tends to S_{max} , ${}^0D_\alpha$ tends to $w_{min}S_{max}$, so ${}^0D_\beta$ tends to $1/w_{min}$. The maximum value β diversity can reach is the inverse of the weight of the smallest community: its contribution to α diversity is proportional to its weight, but its contribution to γ diversity is its richness. Given the weights, the maximum value of β diversity is thus $1/w_{min}$; it is the number of communities if weights are equal.

Comparing β diversity between meta-communities made of different number of communities is not possible without normalization. Jost [3] suggests normalizing it to the unit interval by dividing it by the number of communities in the equal-weight case. We suggest extending this solution to dividing β diversity by $1/w_{min}$. When weights are not equal, the number of communities is not the appropriate reference.

Although we could come back to the equally-weighted-community partition case, β diversity is not independent of α diversity because communities are not independent of each other (some are repeated). Chao *et al.* (appendix B1 of [7]) derive the relation between the maximum value of ${}^0D_\beta$ and ${}^0D_\alpha$ for a two-community meta-community: ${}^0D_\beta \leq \frac{1}{w_{min}} [1 - \frac{w_{max} - w_{min}}{{}^0D_\alpha}]$. The last term quantifies the relation between α and β diversity. It vanishes when weights are close to each other, and it decreases quickly with ${}^0D_\alpha$. If α diversity is not too low (say 50 species), the constraint is negligible (${}^0D_\beta$ can be greater than $0.98/w_{min}$ whatever the weights).

A complete study of the dependence between α and β diversity for all q values and more than two communities is beyond the scope of this paper but these first results show that this dependence is not so serious a problem as that between α and β entropy. As long as weights are not too unequal and diversity is not too small, results can be interpreted clearly.

Very unequal weights imply lower β diversity: the extreme case is when the larger community is the richest. If it is large enough, the meta-community is essentially made of the largest community and ${}^0D_\beta$ tends to 1. This is not an issue of the measure, but a consequence of the sampling design.

Conclusion

The additive framework we proposed here has the advantage of generalizing the widely-accepted decomposition of Shannon entropy, providing a self-contained definition of β entropy and some ways to correct for estimation biases. Deformed logarithms allow a formal parallelism between HCDT and Shannon entropy

(equations (15) and (16) and Table 1). Of course, diversity can be calculated directly, but no estimation-bias correction is available then. The additive decomposition of HCDT entropy can be considered empirically as a calculation tool whose results must systematically be converted to diversity for interpretation.

We rely on Routledge's definition of α entropy which allows decomposing unequally-weighted communities and takes place in a well-established theoretical framework following Patil and Taillie. The price to pay is some dependence between α and β diversity when weights are not equal. It appears to be acceptable since it is unlikely to lead to erroneous conclusions. Still, a rigorous quantifying of it shall be the object of future research.

We only considered communities where individuals were identified and counted, such as forest inventories. Entropy decomposition remains valid when frequencies only are available but our bias correction relies entirely on the number of individual: other techniques will have to be developed for these communities if unobserved species cannot be neglected. Bias correction is still an open question. We proposed a first and rough solution. More research is needed to combine the available approaches rather than using each of them in turn.

References

1. Tuomisto H (2010) A diversity of beta diversities: straightening up a concept gone awry. part 1. defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33: 2–22.
2. Jost L (2006) Entropy and diversity. *Oikos* 113: 363–375.
3. Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88: 2427–2439.
4. Jost L (2008) Gst and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015–4026.
5. Jost L, DeVries P, Walla T, Greeney H, Chao A, et al. (2010) Partitioning diversity for conservation analyses. *Diversity and Distributions* 16: 65–76.
6. Ellison AM (2010) Partitioning diversity. *Ecology* 91: 1962–1963.
7. Chao A, Chiu CH, Hsieh TC (2012) Proposing a resolution to debates on diversity partitioning. *Ecology* 93: 2037–2051.
8. Havrda J, Charvát F (1967) Quantification method of classification processes. concept of structural α -entropy. *Kybernetika* 3: 30–35.
9. Tsallis C (1988) Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics* 52: 479–487.
10. Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5–13.
11. Hill MO (1973) Diversity and evenness: A unifying notation and its consequences. *Ecology* 54: 427–432.
12. Marcon E, Hérault B, Baraloto C, Lang G (2012) The decomposition of shannon's entropy and a confidence interval for beta diversity. *Oikos* 121: 516–522.
13. Borland L, Plastino AR, Tsallis C (1998) Information gain within nonextensive thermostatics. *Journal of Mathematical Physics* 39: 6490–6501.
14. Patil GP, Taillie C (1982) Diversity as a concept and its measurement. *Journal of the American Statistical Association* 77: 548–561.
15. Shannon CE (1948) A mathematical theory of communication. *The Bell System Technical Journal* 27: 379–423, 623–656.
16. Simpson EH (1949) Measurement of diversity. *Nature* 163: 688.
17. Kullback S, Leibler RA (1951) On information and sufficiency. *The Annals of Mathematical Statistics* 22: 79–86.
18. Rao C, Nayak T (1985) Cross entropy, dissimilarity measures, and characterizations of quadratic entropy. *Information Theory, IEEE Transactions on* 31: 589–593.
19. Routledge R (1979) Diversity indices: Which ones are admissible? *Journal of Theoretical Biology* 76: 503–515.
20. Tsallis C, Mendes RS, Plastino AR (1998) The role of constraints within generalized nonextensive statistics. *Physica A* 261: 534–554.
21. Tsallis C (1994) What are the numbers that experiments provide? *Química Nova* 17: 468–471.
22. MacArthur RH (1965) Patterns of species diversity. *Biological Reviews* 40: 510–533.
23. Lin J (1991) Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory* 37: 145–151.
24. Lamberti PW, Majtey AP (2003) Non-logarithmic jensen-shannon divergence. *Physica A: Statistical Mechanics and its Applications* 329: 81–90.
25. Furuichi S, Yanagi K, Kuriyama K (2004) Fundamental properties of tsallis relative entropy. *Journal of Mathematical Physics* 45: 4868–4877.
26. Dauby G, Hardy OJ (2012) Sampled-based estimation of diversity sensu stricto by transforming hurst diversities into effective number of species. *Ecography* 35: 661–672.
27. Chao A, Shen TJ (2003) Nonparametric estimation of shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics* 10: 429–443.
28. Horvitz D, Thompson D (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47: 663–685.
29. Good IJ (1953) On the population frequency of species and the estimation of population parameters. *Biometrika* 40: 237–264.
30. Bonachela JA, Hinrichsen H, Muñoz MA (2008) Entropy estimates of small data sets. *Journal of Physics A: Mathematical and Theoretical* 41: 1–9.
31. Grassberger P (1988) Finite sample corrections to entropy and dimension estimates. *Physics Letters A* 128: 369–373.
32. Fisher RA, Corbet AS, Williams CB (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12: 42–58.
33. Holte D, Große I, Herzel H (1998) Bayes' estimators of generalized entropies. *Journal of Physics A: Mathematical and General* 31: 2551–2566.
34. Hou Y, Wang B, Song D, Cao X, Li W (2012) Quadratic tsallis entropy bias and generalized maximum entropy models. *Computational Intelligence*.
35. Chao A, Wang YT, Jost L (2013) Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods in Ecology and Evolution* 4: 1091–1100.
36. Leinster T, Cobbold C (2011) Measuring diversity: the importance of species similarity. *Ecology* 93: 477–489.
37. Hubbell SP, Condit R, Foster RB (2005) Barro colorado forest census plot data. Available: <https://ctfs.arnarb.harvard.edu/webatlas/datasets/bci>.
38. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. *vegan: Community ecology package*. Available: <http://CRAN.R-project.org/package=vegan>.
39. R Development Core Team (2013) R: A language and environment for statistical computing.
40. Chao A (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11: 265–270.
41. Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70: 3321–3323.

We provide the necessary code for R to compute the analyses presented in this paper as a supplementary material in Appendix S4 with a short user's guide in Appendix S3.

Supporting Information

Appendix S1 Detailed derivation of the partitioning. (PDF)

Appendix S2 Decomposition of Simpson index. (PDF)

Appendix S3 Using the code: short user's guide. (PDF)

Appendix S4 R code to compute the analyses. (ZIP)

Author Contributions

Wrote the paper: EM. Developed the core methods: EM. Contributed to methods: IS BH VR GL. Wrote the R code: EM BH.