

The Rich-Gini-Simpson quadratic index of biodiversity

Radu Cornel Guiasu¹, Silviu Guiasu²

¹Environmental and Health Studies Program, Department of Multidisciplinary Studies, Glendon College, York University, Toronto, Canada; rguiasu@glendon.yorku.ca;

²Department of Mathematics and Statistics, Faculty of Pure and Applied Sciences, York University, Toronto, Canada; guiasus@pascal.math.yorku.ca.

Received 6 August 2010; revised 10 September 2010; accepted 13 September 2010.

ABSTRACT

The Gini-Simpson quadratic index is a classic measure of diversity, widely used by ecologists. As shown recently, however, this index is not suitable for the measurement of beta diversity when the number of species is very large. The objective of this paper is to introduce the Rich-Gini-Simpson quadratic index which preserves all the qualities of the classic Gini-Simpson index but behaves very well even when the number of species is very large. The additive partitioning of species diversity using the Rich-Gini-Simpson quadratic index and an application from island biogeography are analyzed.

Keywords: Rich-Gini-Simpson Index of Species Diversity; Additive Partitioning of Diversity; Island Biogeography; Biodiversity

1. INTRODUCTION

Measuring the diversity of species in a habitat has been an important area of interest in fields such as conservation biology, ecology, and biogeography for the last several decades [1]. Let us assume that there are n species and let:

$$p_i > 0, \quad (i = 1, \dots, n), \quad \sum_i p_i = 1, \quad (1)$$

be the relative frequency distribution of these species in the respective habitat. There are three classic measures of diversity:

- The number of species, or richness: n ;
- The Gini-Simpson quadratic index (abbreviated in this paper as GS):

$$GS = \sum_i p_i(1 - p_i) = 1 - \sum_i p_i^2, \quad (2)$$

introduced by Gini [2] and adapted for biological studies by Simpson [3];

- The Shannon entropy (abbreviated as H):

$$H = -\sum_i p_i \ln p_i, \quad (3)$$

introduced by Shannon [4], as the discrete variant of the continuous entropy defined by Boltzmann [5] in statistical mechanics. There is an extensive literature [1,6-17] about the properties and applications of these measures of diversity.

When the number of species n and the relative abundance of species (1) are the only sources of information available, many other measures of diversity have been proposed. Recently, Jost [18,19] pleaded in favor of the “true” measure of diversity, introduced by Hill [20]:

$$N_r = \left(\sum_i p_i^r \right)^{1/(1-r)}. \quad (4)$$

For $r=0$, we get: $N_0 = n$. For $r=1$, N_r is not defined because the denominator of the exponent, $1-r$, is equal to zero. However, the limit of N_r when r tends to 1 is $\exp(H)$. For $r=2$ we get $1/(1-GS)$. In fact, the natural logarithm of (4), *i.e.* $\ln N_r$, is just Rényi’s entropy [21]. There are no sound reasons to call (4) a “true” measure of diversity. It is simply a unifying notation, as mentioned in [20]. Besides, by performing mathematical transformations on classic measures of diversity, like taking the exponential of the Shannon entropy or the reciprocal of the Gini-Simpson quadratic index, for example, we obtain other measures that lose, however, some essential features of the original measures, such as concavity, for instance. Concavity is an essential property of any measure that can be used in an additive partitioning of species diversity. Hoffmann and Hoffmann [22] are right when asking: “Is there a ‘true’ measure of diversity?” As noticed by Ricotta [23], there is a “jungle of measures of diversity” in the current conservation biology literature. Under the circumstances, perhaps the best strategy is to remember Occam’s razor and, trying to keep things simple, it may be easier to just go

back to the classic measures of diversity mentioned before and see how they can be adjusted to address new problems under new circumstances.

The species richness n is very simple but ignores the abundance of species. Shannon's entropy has excellent properties but is difficult to estimate and maximizing it subject to linear constraints, generally gives a solution satisfying exponential equations which cannot be solved analytically. On the other hand, the Gini-Simpson quadratic index is simpler and generally seems to be preferred by ecologists. Jost [18,19,24], however, noticed a troubling anomaly related to GS . Indeed, if this measure is used in the additive partitioning of species diversity, the corresponding beta diversity approaches zero when the number of species is very large. Thus, for two habitats with no species in common, for instance, the between-habitat diversity tends to zero when the number of species in one of the habitats, or in both of them, tends to infinity, instead of becoming larger as is obviously the case in actual fact.

The objective of this paper is to show that the anomaly just mentioned can be easily fixed. The product between the species richness a) and the measure of diversity c), called here the Rich-Gini-Simpson quadratic index and abbreviated as RGS , preserves all the basic properties of GS and behaves well when the number of species is large. Therefore, RGS is suitable for use in the additive partitioning of species diversity. Subsequently, the RGS index is applied to data on the avifaunal diversity on several tropical Indian Ocean islands, using some of the numerical results obtained by Adler [25], in order to show how the alpha, beta, and gamma species diversities change when the usual equal weights for the various habitats are replaced by the relative areas and the relative elevations of the respective islands.

2. THE RICH-GINI-SIMPSON INDEX

If there are n species in a certain habitat and their relative abundance is given by (1), the Rich-Gini-Simpson quadratic index is

$$RGS = n \sum_i p_i (1 - p_i) = n(1 - \sum_i p_i^2). \quad (5)$$

The concavity of RGS and the maximum value of RGS are analyzed in the Appendix. Thus, we have: $0 \leq RGS \leq n-1$, the maximum corresponding to the uniform distribution: $p_i = 1/n$, ($i = 1, \dots, n$). As $GS = RGS/n$, the maximum value of GS is $\max_p GS = 1 - 1/n$, corresponding to the uniform distribution as well. The essential difference between these two indexes is that GS is bounded by 1 and tends to 1 when the number of species n tends to infinity,

Identify applicable sponsor/s here. (*sponsors*)

whereas RGS is not bounded and tends to infinity when the number of species n tends to infinity. Shannon's entropy, on the other hand, has the maximum $\max_p H = \ln n$, which tends to infinity when the number of species n tends to infinity, but it increases much more slowly than $\max_p RGS = n - 1$.

Pleading against the use of the GS index, Jost [24] gave the following example: "Suppose a continent has a million equally-common species, and a meteor impact kills 999,900 of the species, leaving 100 species untouched. Any biologist, if asked, would say that this meteor impact caused a large absolute and relative drop in diversity. Yet GS only decreases from 0.999999 to 0.99, a drop of less than 1%". Jost concluded that: "[The ecologists relying on GS will often misjudge the magnitude of ecosystem changes. This same problem arises when Shannon entropy is equated with diversity. In contrast, $N_2 \dots$ drops by the intuitively appropriate 99.99%". This example shows that there is indeed a troubling anomaly in using GS when the number of species is very large. But RGS has no such drawback. Indeed, if before the cataclysm there are $n = 1,000,000$ equally abundant species, then:

$$GS = 1 - \frac{1}{n} = 0.999999; \quad H = \ln n = 13.8155105;$$

$$\exp(H) = \exp(\ln n) = n = 1000000,$$

$$N_2 = n = 1000000, \quad RGS = n - 1 = 999999.$$

After the cataclysm, there are only $n = 100$ equally abundant species left. Thus:

$$GS = 1 - \frac{1}{n} = 0.99; \quad H = 4.605170186;$$

$$\exp(H) = \exp(\ln n) = n = 100,$$

$$N_2 = n = 100, \quad RGS = n - 1 = 99.$$

Therefore, GS indicates a decrease in diversity equal to 0.999901%, which is obviously wrong, H indicates a decrease in diversity equal to 66.66666667%, which is not good enough, whereas RGS , $\exp(H)$ and N_2 give a decrease in diversity equal to 99.99009999% and 99.99000000%, respectively, in agreement with common sense. Let us note that, practically, RGS and $\exp(H)$ have the same maximum value when the number of species n is given, but the index $\exp(H)$ is not a concave function of the relative frequency distribution of species $p = (p_1, \dots, p_n)$ and, consequently, it is not suitable to be used in the additive partitioning of species diversity, whereas the index RGS is.

3. THE ADDITIVE PARTITIONING OF SPECIES DIVERSITY USING RGS

MacArthur [26] pointed out the need for a theory of within-habitat and between-habitat species diversities.

He, together with Recher and Cody [27], proposed a measure of the difference between the species diversities of two habitats based on Shannon's entropy and using equal relative weights for habitats. This measure was also used in the influential book "The theory of island biogeography" by MacArthur and Wilson [28]. Rao [29], without mentioning the paper [27], extended the measure of the difference between the species diversities of two habitats, based also on Shannon's entropy but using arbitrary weights assigned to the two habitats. Whittaker [30,31] proposed linking diversity components between ecological scales by multiplication so that the gamma diversity, measuring the species diversity in a larger region consisting of several ecological communities taken together, is the product of the alpha diversity, which measures the mean species diversity in the local communities taken separately, and the beta diversity, representing the variation and changes in mean species diversity in a larger region which contains the local ecological communities taken together, as a whole. The beta diversity essentially measures the biogeographic changes in species diversity among various locations within a larger region. As such, the beta diversity can be important in leading to the development of geographic strategies for the conservation of species and habitats, as mentioned by Harrison and Quinn [32]. Routledge [33,34] developed Whittaker's approach. Allan [35] applied an additive linkage of species diversity components according to which the gamma diversity is partitioned into the sum of the alpha diversity and the beta diversity, using the Shannon entropy. Lande [36] dealt with an arbitrary number of habitats and arbitrary weights, using the Shannon entropy, and extended this approach to species richness and to the Gini-Simpson index, recommending the additive partitioning of species diversity as a unifying framework for measuring species diversity at different levels of ecological organization. As mentioned by Wagner, Wildi and Ewald [37], in contrast to the multiplicative model, by using the additive partitioning, all species diversity components are measured in the same way and expressed in the same units, so that they can be directly compared. Recently, it was pointed out that the additive partitioning of species diversity is an old idea which shows a new revival. According to Veech, Summerville, Crist and Gering [38], "Lande [36] appears to have been the first to place the additive partitioning of species diversity in the context of Whittaker's concepts of alpha, beta, and gamma diversities ... Viewing gamma diversity as the sum of alpha and beta diversities leads to the most operational definition of beta diversity and quantifies it in a manner comensurate with the measurement of alpha and gamma diversities. In effect, the revival of

additive diversity partitioning has given new meaning to beta diversity".

As RGS is a concave function, it is suitable for the additive partitioning of species diversity. Let $\{x_1, \dots, x_n\}$ be a set of species and let $\{x_i, i \in I\}$ and $\{x_i, i \in J\}$ be the species from the habitats h_1 and h_2 , respectively. The number of species from h_1 is n_1 and the number of species from h_2 is n_2 . Obviously, $n_1 \leq n$, $n_2 \leq n$, and $n \leq n_1 + n_2$. The species $\{x_i, i \in I - J\}$ belong only to the habitat h_1 , the species $\{x_i, i \in J - I\}$ belong only to the habitat h_2 , whereas the species $\{x_i, i \in I \cap J\}$ belong to both habitats. We have $I \cup J = \{1, \dots, n\}$.

Let $\{p_i, i \in I\}$ and $\{q_i, i \in J\}$ be the relative frequencies of the species from h_1 and h_2 , respectively. We have:

$$p_i > 0, \quad \sum_{i \in I} p_i = 1; \quad q_i > 0, \quad \sum_{i \in J} q_i = 1.$$

In general, the beta diversity is the average between-habitat diversity, whereas the alpha diversity is the average diversity of the individual communities or the average within-habitat diversity. Using the additive partitioning of the species diversity, the gamma diversity is the sum of the alpha and beta diversities, or the average total diversity. Let $\lambda_1 > 0$, and $\lambda_2 > 0$, be two weights assigned to the habitats h_1 and h_2 , respectively, such that $\lambda_1 + \lambda_2 = 1$. We use these weights to calculate the average within-habitat species diversity, *i.e.* the alpha diversity, and the average relative frequency of the species used in the total species diversity of a larger region that includes the two individual habitats, *i.e.* the gamma diversity. If the two weights are equal, namely $\lambda_1 = \lambda_2 = 1/2$, then the average is just the arithmetic mean. These weights, however, may represent the relative areas or the relative elevation of the two habitats, or any other quantitative characteristics of the habitats that can affect the diversity of the species. In this context, alpha diversity refers to the average species diversity in the two habitats h_1 and h_2 , taken separately, gamma diversity refers to the species diversity in the habitats h_1 and h_2 , averaged together, whereas beta diversity represents the average between-habitat species diversity as we move from the individual habitats h_1 , h_2 , averaged separately, to the larger region containing the union of h_1 and h_2 , averaged together. We now use RGS to calculate the alpha, gamma, and beta species diversities. Denote by:

$$RGS(h_1) = n_1 \sum_{i \in I} p_i (1 - p_i) = n_1 \left(1 - \sum_{i \in I} p_i^2 \right),$$

$$RGS(h_2) = n_2 \sum_{i \in J} q_i (1 - q_i) = n_2 \left(1 - \sum_{i \in J} q_i^2 \right),$$

in which case the alpha diversity is:

$$\alpha - Div = \lambda_1 RGS(h_1) + \lambda_2 RGS(h_2).$$

The gamma-diversity is:

$$\begin{aligned} \gamma - Div &= RGS(\lambda_1, h_1; \lambda_2, h_2) = \\ &= n \sum_{i \in I \cup J} (\lambda_1 p_i + \lambda_2 q_i)(1 - \lambda_1 p_i - \lambda_2 q_i) = \\ &= n \left[1 - \sum_{i \in I \cup J} (\lambda_1 p_i + \lambda_2 q_i)^2 \right], \end{aligned}$$

where $p_i = 0$ for $i \in J - I$, and $q_i = 0$ for $i \in I - J$.

The concavity of RGS allows the additive partition of species diversity, and the beta-diversity is:

$$\begin{aligned} \beta - Div &= (\gamma - Div) - (\alpha - Div) = \\ &= (n - \lambda_1 n_1 - \lambda_2 n_2) + \lambda_1 (n_1 - \lambda_1 n) \sum_{i \in I} p_i^2 + \\ &+ \lambda_2 (n_2 - \lambda_2 n) \sum_{i \in J} q_i^2 - 2 \lambda_1 \lambda_2 n \sum_{i \in I \cap J} p_i q_i \geq 0. \end{aligned}$$

If both habitats contain the same species, then $I = J$, which implies $I \cap J = I$, $I \cup J = I$, $n_1 = n_2 = n$, and the beta-diversity has a simple expression:

$$\begin{aligned} \beta - Div &= \lambda_1 \lambda_2 n \left(\sum_{i \in I} p_i^2 + \sum_{i \in I} q_i^2 - 2 \sum_{i \in I} p_i q_i \right) = \\ &= \lambda_1 \lambda_2 n \sum_{i \in I} (p_i - q_i)^2. \end{aligned}$$

Clearly, if both habitats have the same species and the same abundance of these species, namely $p_i = q_i$, ($i = 1, \dots, n$), then $\beta - Div = 0$.

If the two habitats have no species in common, then $I \cap J = \emptyset$, $n = n_1 + n_2$, and the beta-diversity is:

$$\begin{aligned} \beta - Div &= (n - \lambda_1 n_1 - \lambda_2 n_2) + \\ &+ \lambda_1 (n_1 - \lambda_1 n) \sum_{i \in I} p_i^2 + \lambda_2 (n_2 - \lambda_2 n) \sum_{i \in J} q_i^2. \end{aligned}$$

In particular, if the two habitats have no species in common and in each habitat the species have the same abundance, namely $p_i = 1/n_1$, ($i \in I$), and $q_i = 1/n_2$, ($i \in J$), then the beta-diversity is:

$$\begin{aligned} \beta - Div &= \\ &= \lambda_1 n_2 \left(1 - \sum_{i \in I} p_i^2 \right) + \lambda_2 n_1 \left(1 - \sum_{i \in J} q_i^2 \right) = \\ &= \lambda_1 n_2 \left(1 - \frac{1}{n_1} \right) + \lambda_2 n_1 \left(1 - \frac{1}{n_2} \right), \end{aligned}$$

which tends to $+\infty$ if n_1 tends to $+\infty$ or 1 and n_2 tends to $+\infty$.

Remark 1. The generalization of the results from this section to the case of an arbitrary number of habitats h_1, \dots, h_m is straightforward.

Remark 2. As mentioned by Lande [36], the ratio between the alpha diversity and the gamma diversity may be used as a similarity index, denoted here by Sim .

Arguing against the use of the GS index and the additive partitioning of species diversity, Jost [19] discussed

the following example: “Suppose a continent with 30 million equally common species is hit by a plague that kills half the species. How do some popular diversity indices judge this drop in diversity? ... The Shannon entropy only drops from 17.2 to 16.5; according to this index the plague caused a drop of only 4% in the ‘diversity’ of the continent. This does not agree well with our intuition that the loss of half the species and half the individuals is a large drop in diversity. The Gini-Simpson index drops from 0.99999997 to 0.99999993; if this index is equated with ‘diversity’, the continent has lost practically no ‘diversity’ when half its species and individuals disappeared”. Instead of GS and H , Jost proposes the use of $\exp(H)$, which in his example has the value:

$$\exp(H) = \exp(\ln 30000000) = 30000000$$

before the plague and:

$$\exp(H) = \exp(\ln 15000000) = 15000000$$

after the plague, corresponding to a loss of 50% in diversity. However, as $\exp(H)$ is not a concave function, the additive partitioning of species diversity cannot be used and should be replaced by the multiplicative partitioning of species diversity as Whittaker [30,31] and Routledge [33,34] proposed. The situation, however, is not as hopeless as it may seem to be. In fact, it is not really hopeless at all. The additive partitioning of species diversity, so popular with some ecologists because it allows the alpha, beta, and gamma diversities to be measured in the same way and be expressed in the same units so that they can be directly compared, may in fact be preserved but GS has to be replaced by RGS . Thus, in the case just mentioned:

$$RGS = n - 1 = 30000000 - 1 = 29999999$$

before the plague and:

$$RGS = n - 1 = 15000000 - 1 = 14999999$$

after the plague, corresponding to a loss of 50% in diversity, in total agreement with common sense.

Example: If there are 30,000,000 species uniformly distributed in habitat h_1 and 15,000,000 of these species are uniformly distributed in habitat h_2 , then, using the equal weights $\lambda_1 = \lambda_2 = 1/2$ and the GS index, we obtain:

$$GS(h_1) = 1 - \frac{1}{30000000} = 0.99999997,$$

$$GS(h_2) = 1 - \frac{1}{15000000} = 0.99999993,$$

which show almost no difference in species diversities. Also:

$$\alpha - Div = \frac{1}{2} GS(h_1) + \frac{1}{2} GS(h_2) = 0.99999995,$$

$$\gamma - Div = GS(\lambda_1, h_1; \lambda_2, h_2) =$$

$$= 1 - 15000000 \times \left(\frac{1}{60000000} \right)^2 - 15000000 \times \left(\frac{3}{60000000} \right)^2 = 0.9999999,$$

$$\beta - Div = (\gamma - Div) - (\alpha - Div) = 0.000000083,$$

which shows that the between-habitat species diversity is practically zero, in contrast to the fact that h_1 has a much higher species diversity than h_2 . The similarity index is:

$$Sim = \frac{0.9999999}{0.99999995} = 0.99999995.$$

Using the Shannon entropy:

$$H(h_1) = \ln 30000000 = 17.2167,$$

$$H(h_2) = \ln 15000000 = 16.5235,$$

which show a very small difference in diversity, in fact a decrease of only 4.03% in h_2 with respect to h_1 , contrary to common sense. Also:

$$\alpha - Div = \frac{1}{2}H(h_1) + \frac{1}{2}H(h_2) = 16.8702,$$

$$\gamma - Div = H(\lambda_1, h_1; \lambda_2, h_2) =$$

$$= 15000000 \times \left(-\frac{1}{60000000} \ln \frac{1}{60000000} \right) +$$

$$+ 15000000 \times \left(-\frac{3}{60000000} \ln \frac{3}{60000000} \right) = 17.0859,$$

$$\beta - Div = (\gamma - Div) - (\alpha - Div) = 0.2157,$$

a very small between-habitat species diversity. The similarity index is:

$$Sim = \frac{16.8702}{17.0859} = 0.98737756,$$

which is much too high.

Using now the equal weights $\lambda_1 = \lambda_2 = 1/2$ and the RGS diversity index, we obtain:

$$RGS(h_1) = 30000000 - 1 = 29999999,$$

$$RGS(h_2) = 15000000 - 1 = 14999999,$$

showing a decrease of 50% in species diversity in h_2 compared to h_1 , in complete agreement with common sense. Also:

$$\alpha - Div = \frac{1}{2}RGS(h_1) + \frac{1}{2}RGS(h_2) = 2.25 \times 10^7,$$

$$\gamma - Div = RGS(\lambda_1, h_1; \lambda_2, h_2) =$$

$$= 30000000 \times \left[1 - 15000000 \times \left(\frac{1}{60000000} \right)^2 - \right.$$

$$\left. - 15000000 \times \left(\frac{3}{60000000} \right)^2 \right] = 3 \times 10^7,$$

$$\beta - Div = (\gamma - Div) - (\alpha - Div) = 7.5 \times 10^6,$$

which show that the average between-habitat species diversity is 25% of the average total species diversity, whereas the average within-habitat species diversity is 75% of the average total species diversity. However, there are similarities between the two habitats, in the sense that h_2 contains half of the species of h_1 , there are no species from h_2 that are not found in h_1 , and both h_1 and h_2 have their species uniformly distributed. These features make h_2 somewhat similar to h_1 . Using RGS, the similarity index is:

$$Sim = \frac{2.25 \times 10^7}{3 \times 10^7} = 0.75.$$

Remark 3. If habitat h_1 contains only one species x_1 and habitat h_2 contains only one species x_2 , then, obviously:

$$RGS(h_1) = 0, \quad RGS(h_2) = 0, \quad \alpha - Div = 0,$$

$$\gamma - Div = RGS\left(\frac{1}{2}, h_1; \frac{1}{2}, h_2\right) = 2 \left(1 - \frac{1}{4} - \frac{1}{4} \right) = 1,$$

$$\beta - Div = 1, \quad Sim = \frac{0}{1} = 0.$$

4. APPLICATION

There are many discussions of the role and applications of the measures of species diversity in biogeography (for instance, [15,35,39-43]). For example, MacArthur and Wilson [28] analyzed the impact of factors such as island area and the distance between the island and the mainland on the species diversity found on various islands. Some of the findings of this classic study were also applied to the study of habitat islands and nature reserves, as well as real islands, surrounded by the sea [43-45]. When MacArthur, Recher and Cody [27] introduced their measure of the average difference in species diversity between two habitats, they assigned equal weights to the respective habitats, taking into account only the relative frequencies of the species from the two habitats. More often than not, however, the habitats could be very different in other respects, and some additional factors, like area or elevation, for instance, may also have to be taken into account even when the habitats are located in the same general geographic region. These factors may be given various weights, which can be taken into account when calculating the alpha, beta, and gamma species diversities. If there are two habitats h_1 , h_2 , and their areas (in km^2) are a_1 and a_2 , respectively, then we may attach to the two habitats the weights: $\lambda_1 = a_1/(a_1 + a_2)$ and $\lambda_2 = a_2/(a_1 + a_2)$ respectively. The same approach can be applied if the elevation (or some other factor of interest) is taken into account.

Adler [25] analyzed the birdspecies diversity on 14

different tropical archipelagoes and isolated islands in the Indian Ocean. The 139 species of resident birds, belonging to 33 families, found on these islands were grouped into three main categories: Continental, Indian Ocean (species found only on Indian Ocean islands, in general), and Endemic (species found only on a single Indian Ocean archipelago or island). **Table 1** contains the initial data set consisting of: the absolute frequencies of the Continental species (Cont), Indian Ocean species (IndOc), and Endemic species (End), the area (in km^2), and the elevation (highest peak in m), for seven archipelago / island habitats from the Indian Ocean, as given by Adler [25]. The seven archipelagoes or isolated islands (equivalent to seven distinct habitats for the purposes of this study) are: h_1 : Christmas Island; h_2 : Rodriguez; h_3 : Mauritius; h_4 : Reunion; h_5 : Seychelles; h_6 : Aldabra Islands; h_7 : Comoro Islands. Our objective here is to calculate the numerical values of the alpha, gamma, and beta diversities, using the quadratic index RGS , when the weights assigned to the archipelagoes / islands are equal, or are the relative areas or the relative elevations of the respective archipelagoes / islands.

Table 2 contains: $p_{1,j}$, the relative frequency of Continental species in habitat h_j ; $p_{2,j}$, the relative frequency of Indian Ocean species in habitat h_j ; $p_{3,j}$, the relative frequency of Endemic species in habitat h_j ;

Table 1. Application: The data set.

h_j	Cont	IndOc	End	Area(km^2)	Elevation(m)
h_1	7	0	2	135	361
h_2	1	0	12	119	396
h_3	7	6	15	1865	828
h_4	6	6	15	2512	3069
h_5	7	1	11	258	905
h_6	19	3	1	172	24
h_7	32	4	13	2236	2360

Table 2. Relative frequency and the RGS index.

h_j	$p_{1,j}$	$p_{2,j}$	$p_{3,j}$	$RGS(h_j)$
h_1	0.777778	0.000000	0.222222	0.691358
h_2	0.076923	0.000000	0.923077	0.284024
h_3	0.250000	0.214286	0.535714	1.813776
h_4	0.222222	0.222222	0.555556	1.777779
h_5	0.368421	0.052632	0.578947	1.578948
h_6	0.826087	0.130435	0.043478	0.896031
h_7	0.653061	0.081633	0.265306	1.489380

the RGS index of habitat h_j . We can see that Mauritius has a greater bird species diversity ($RGS = 1.813776$) than the other archipelagoes or islands considered here, followed by Reunion ($RGS = 1.777779$) and Seychelles ($RGS = 1.578948$). The lowest bird species diversity by far is on Rodriguez ($RGS = 0.284024$). These values have to be compared with the maximum value of RGS , which in this application is $n-1 = 3-1 = 2$.

Dealing with seven habitats, we calculate the alpha, gamma, and beta diversities according to the formulas:

$$\alpha - Div = \sum_{j=1}^7 \lambda_j RGS(h_j),$$

$$\gamma - Div = RGS(\lambda_1, h_1; \dots; \lambda_7, h_7) =$$

$$= 3 \times \left[1 - \sum_{i=1}^3 \left(\sum_{j=1}^7 \lambda_j p_{i,j} \right)^2 \right],$$

$$\beta - Div = (\gamma - Div) - (\alpha - Div),$$

where the weights are:

$$\lambda_j > 0, \quad (j=1, \dots, 7), \quad \sum_{j=1}^7 \lambda_j = 1.$$

The similarity index is:

$$Sim = \frac{\alpha - Div}{\gamma - Div}.$$

Case 1. If we take all seven archipelago/island habitats together, as a group, and the weights are:

$$\lambda_j = \frac{area(h_j)}{area(h_1) + \dots + area(h_7)}, (j=1, \dots, 7),$$

we get the corresponding relative area weights:

$$\lambda_1 = 0.018501, \lambda_2 = 0.016308, \lambda_3 = 0.255584,$$

$$\lambda_4 = 0.344251, \lambda_5 = 0.035357, \lambda_6 = 0.023571,$$

$$\lambda_7 = 0.306427,$$

for which we obtain:

$$\alpha - Div = 1.62633, \quad \gamma - Div = 1.86103,$$

$$\beta - Div = 0.234693, \quad Sim = 0.873887.$$

Case 2. If we take all seven archipelago/island habitats together, as a group, and the weights are:

$$\lambda_j = \frac{elevat(h_j)}{elevat(h_1) + \dots + elevat(h_7)}, (j=1, \dots, 7),$$

we get the following relative elevation weights:

$$\lambda_1 = 0.045449, \lambda_2 = 0.049855, \lambda_3 = 0.104243,$$

$$\lambda_4 = 0.386378, \lambda_5 = 0.113937, \lambda_6 = 0.003022,$$

$$\lambda_7 = 0.297117,$$

for which we obtain:

$$\alpha - Div = 1.54668, \quad \gamma - Div = 1.81972,$$

$$\beta - Div = 0.273039, \quad Sim = 0.849955.$$

Case 3. If we take all seven archipelago / island habitats together, as a group, and the weights are equal:

$$\lambda_j = \frac{1}{7}, \quad (j=1, \dots, 7),$$

we obtain the average values:

$$\alpha - Div = 1.21876, \quad \gamma - Div = 1.75528,$$

$$\beta - Div = 0.536528, \quad Sim = 0.694339.$$

Generally, for islands or habitat islands found in a similar geographic region, species diversity tends to be greater on the island or habitat island with a larger area or a higher elevation. The above numerical results obtained by using *RGS* as the main mathematical tool, show that by taking the area and elevation into account, in this order, the alpha and gamma species diversities increase whereas the beta species diversity decreases compared to what happens when we calculate the mean within-habitat and between-habitat species diversity ignoring such factors. Calculating the alpha, beta, and gamma species diversities by using the relative areas and the relative elevation as weights, we compensate for the lack of homogeneity of the habitats with respect to such essential factors which influence species diversity.

5. CONCLUSIONS

The Gini-Simpson index for species diversity is very popular with many ecologists. Recently, however, Jost [18,19,24] showed that this index does not behave well when the number of species is large and is not suitable for use in the computation of the between-habitat species diversity, also called the beta diversity. As a result, Jost pleaded in favour of abandoning the Gini-Simpson index and replacing the additive partitioning of species diversity, preferred by many ecologists, with the multiplicative partitioning. The objective of this paper is to show that the additive partitioning of species diversity may be preserved but the classic Gini-Simpson index of diversity should be replaced by the Rich-Gini-Simpson index, abbreviated as *RGS*, which behaves well when the number of species is large, while keeping the useful basic properties of the classic Gini-Simpson index unchanged. The properties of the *RGS* index and its use in the additive partitioning of the species diversity are analyzed. *RGS* is also applied to data on the avifaunal diversity on several tropical Indian Ocean islands (using some of the numerical data obtained by Adler [25]). The application shows that by using the *RGS* index as a mathematical tool and introducing weights directly proportional with the areas or elevation of the habitats (in this order), the within-habitat species diversity and the total species diversity increase while the between-habitat species diversity decreases compared to what happens when we calculate the mean within-habitat and between-habitat species diversities ignoring such im-

portant factors.

REFERENCES

- [1] Sarkar, S. (2007) From ecological diversity to biodiversity. In: Hull, D.L. and Ruse, M. Eds., *The Philosophy of Biology*, Cambridge University Press, Cambridge, 388-409.
- [2] Gini, C. (1912) Variabilità e mutabilità. In: Pizetti, E. and Salvemini, T. Eds., *Rome: Libreria Eredi Virgilio Veschi*, Memorie di metodologica statistica.
- [3] Simpson, E.H. (1949) Measurement of diversity. *Nature*, **163**, 688.
- [4] Shannon, C.E. (1948) A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379-423, 623-656.
- [5] Boltzmann, L. (1896) Vorlesungen über Gastheorie. J.A. Barth edition, Akademische Druck-u, Leipzig.
- [6] Pielou, E.C. (1966) Shannon's formula as a measure of specific diversity: Its use and misuse. *American Naturalist*, **100**, 463-465.
- [7] Pielou, E.C. (1966) Species-diversity and pattern-diversity in the study of ecological succession. *Journal of Theoretical Biology*, **10**, 370-383.
- [8] Pielou, E.C. (1975) Ecological Diversity. Wiley, New York.
- [9] Guiasu, S. (1977) Information Theory with Applications. McGraw-Hill, New York.
- [10] Patil, G.P. and Taillie, C. (1982) Diversity as a concept and its measurement. *Journal of American Statistical Association*, **77**, 548-567.
- [11] Harper, J.L. and Hawksworth, D.L. (1994) Biodiversity: Measurement and estimation. *Philosophical Transactions: Biological Sciences*, **345**, 5-12.
- [12] Magnussen, S. and Boyle, T.J.B. (1995) Estimating sample size for inference about the Shannon-Weaver and the Simpson indices of species diversity. *Forest Ecology and Management*, **78**, 71-84.
- [13] Krebs, C.J. (2001) Ecology: The experimental analysis of distribution and abundance. 5th Edition, Benjamin Cummings, San Francisco.
- [14] Guiasu, R.C. and Guiasu, S. (2003) Conditional and weighted measures of ecological diversity. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **11**, 283-300.
- [15] Guiasu, R.C. and Guiasu, S. (2003) Entropy in ecology and ethology. Nova Science Publishers, New York.
- [16] Magurran, A.E. (2004) Measuring biological diversity. Blackwell Publishing, Malden MA.
- [17] Maclaurin, J. and Sterelny, K. (2008) What is biodiversity? University of Chicago Press, Chicago and London.
- [18] Jost, L. (2006) Entropy and diversity. *Oikos*, **113**, 363-375.
- [19] Jost, L. (2007) Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427-2439.
- [20] Hill, M. (1973) Diversity and evenness. A unifying notation and its consequences. *Ecology*, **54**, 427-432.
- [21] Rényi, A. (1961) On measures of entropy and information. In: Neyman, J. Ed., *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, 547-561.

- [22] Hoffmann, S. and Hoffmann, A. (2008) Is there a “true” diversity? *Ecological Economics*, **65**, 213-215.
- [23] Ricotta, C. (2005) Through the jungle of biological diversity. *Acta Biotheoretica*, **53**, 29-38.
- [24] Jost, L. (2009) Mismeasuring biological diversity: Response to Hoffmann and Hoffmann. *Ecological Economics*, **68**, 925-928.
- [25] Adler, G.H. (1994) Avifaunal diversity and endemism on tropical Indian Ocean islands. *Journal of Biogeography*, **21**, 85-95.
- [26] MacArthur, R.H. (1965) Patterns of species diversity. *Biological Review*, **40**, 510-533.
- [27] MacArthur, R.H., Recher, H. and Cody, M. (1966) On the relation between habitat selection and species diversity. *American Naturalist*, **100**, 319-332.
- [28] MacArthur, R.H. and Wilson, E.O. (1967) The theory of island biogeography. Princeton University Press, Princeton.
- [29] Rao, C.R. (1982) Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, **21**, 24-43.
- [30] Whittaker, R.H. (1972) Evolution and measurement of species diversity. *Taxon*, **21**, 213-251.
- [31] Whittaker, R.H. (1977) Evolution of species diversity in land communities. In: Hecht, M.K. and Steere, B.W.N.C. Eds., *Evolutionary Biology*, Plenum Press, New York, **10**, 1-67.
- [32] Harrison, S. and Quinn, J. (2006) The importance of β - diversity. In: Groom, M.J., Meffe, G.K. and Carroll, C.R. Eds., *Principles of Conservation Biology*, 3rd Edition, Sinauer Associates Inc. Publishers, Sunderland, 44-45.
- [33] Routledge, R.D. (1977) On Whittaker’s components of diversity. *Ecology*, **58**, 1120-1127.
- [34] Routledge, R.D. (1979) Niche metrics and diversity components. *Oecologia*, **43**, 121-124.
- [35] Allan, J.D. (1975) Components of diversity. *Oecologia*, **18**, 359-367.
- [36] Lande, R. (1996) Statistics and partitioning of species diversity and similarity among multiple communities. *Oikos*, **76**, 5-13.
- [37] Wagner, H.H., Wildi, O. and Ewald, K.C. (2000) Additive partitioning of plant species diversity in an agricultural mosaic landscape. *Landscape Ecology*, **15**, 219-227.
- [38] Veech, J.A., Summerville, K.S., Crist, T.O. and Gering, J.C. (2002) The additive partitioning of species diversity: Recent revival of an old idea. *Oikos*, **99**, 3-9.
- [39] Tramer, E.J. (1969) Bird species diversity: Components of Shannon’s formula. *Ecology*, **50**, 927-929.
- [40] Fager, E.W. (1972) Diversity: A sampling study. *American Naturalist*, **106**, 293-310.
- [41] Brown, J.H. and Lomolino, M.V. (1998) *Biogeography*. 2nd Edition, Sinauer Associates Inc. Publishers, Sunderland.
- [42] Stirling, G. and Wilsey, B. (2001) Empirical relationships between species richness, evenness, and proportional diversity. *American Naturalist*, **158**, 286-299.
- [43] Cox, C.B. and Moore, P.D. (2005) *Biogeography: An ecological and evolutionary approach*. 7th Edition, Blackwell Publishing, Oxford.
- [44] Quammen, D. (1997) The song of the Dodo-Island biogeography in an Age of Extinction. Simon & Schuster Inc., New York.
- [45] Powledge, F. (2003) Island biogeography’s lasting impact. *BioScience*, **53**, 1032-1038.