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; guisasus@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:

\[
p_i > 0, \quad (i = 1, \ldots, n), \quad \sum p_i = 1,
\]

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

a) The number of species, or richness: \( n \);

b) The Gini-Simpson quadratic index (abbreviated in this paper as \( GS \)):

\[
GS = \sum p_i (1 - p_i) = 1 - \sum p_i^2,
\]

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

c) The Shannon entropy (abbreviated as \( H \)):

\[
H = -\sum p_i \ln p_i,
\]

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 \( p_i \) 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]:

\[
N_r = \left( \sum p_i \right)^{1/(1 - r)}.
\]

For \( r = 0 \), we get: \( N_0 = n \). For \( r = 1 \), \( N_1 \) 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 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

\[
RGS = n \sum_i p_i (1 - p_i) = n(1 - \sum_i p_i^2). \tag{5}
\]

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: \( p_i = 1/n \), \( i = 1, \ldots, n \). As \( GS = RGS/n \), the maximum value of \( GS \) is maximized, \( GS = 1 - 1/n \), 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, 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 \( \max_n H = \ln n \), which tends to infinity when the number of species \( n \) tends to infinity, but it increases much much more slowly than \( \max_n RGS = n - 1 \).

Pleasing 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, \( N_2 \)... 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 a 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 66.6666667%, which is not good enough, whereas \( RGS \), \( \exp(H) \) and \( N_2 \) give a decrease in diversity equal to 99.9990000999% and 99.99900000 0%, 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 \( p = (p_1, \ldots, p_n) \) 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.

Identify applicable sponsor/s here. (sponsors)
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 commensurate 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 $\{x_1, \ldots, x_n\}$ be a set of species and let $\{x_i, i \in I\}$ and $\{x_i, i \in J\}$ be the species from the habitats $h_i$ and $h_i$, respectively. The number of species from $h_i$ is $n_i$ and the number of species from $h_i$ is $n_i$. Obviously, $n_i \leq n$, $n_i \leq n$, and $n \leq n_i + n_i$. The species $\{x_i, i \in I - J\}$ belong only to the habitat $h_i$, the species $\{x_i, i \in J - I\}$ belong only to the habitat $h_i$, whereas the species $\{x_i, i \in I \cap J\}$ belong to both habitats. We have $I \cup J = \{1, \ldots, n\}$.

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

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

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 $\lambda_i > 0$, and $\lambda_i > 0$, be two weights assigned to the habitats $h_i$ and $h_i$, respectively, such that $\lambda_i + \lambda_i = 1$. 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 $\lambda_i = \lambda_i = 1/2$, 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 $h_i$ and $h_i$, taken separately, gamma diversity refers to the species diversity in the habitats $h_i$ and $h_i$, averaged together, whereas beta diversity represents the average between-habitat species diversity as we move from the individual habitats $h_i$, $h_i$, averaged separately, to the larger region containing the union of $h_i$ and $h_i$, averaged together. We now use $RGS$ to calculate the alpha, gamma, and beta species diversities. Denote by:

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

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

in which case the alpha diversity is:

Copyright © 2010 SciRes.
\[ \alpha - \text{Div} = \lambda_1 \text{RGS}(h_1) + \lambda_2 \text{RGS}(h_2). \]

The gamma-diversity is:
\[ \gamma - \text{Div} = \text{RGS}(\lambda_1, h_1; \lambda_2, h_2) = \]
\[ = n \sum_{i \in I \cup J} (\lambda_1 p_i + \lambda_2 q_i)(1 - \lambda_1 p_i - \lambda_2 q_i) = \]
\[ = n \left[ 1 - \sum_{i \in I \cup J} (\lambda_1 p_i + \lambda_2 q_i)^2 \right]. \]

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

The concavity of \( RGS \) allows the additive partitioning of species diversity, and the beta-diversity is:
\[ \beta - \text{Div} = (\gamma - \text{Div}) - (\alpha - \text{Div}) = \]
\[ = (n - \lambda_1 n_1 - \lambda_2 n_2) + \lambda_1 (n_1 - \lambda_1 n) \sum_{i \in I} q_i^2 + \]
\[ + \lambda_2 (n_2 - \lambda_2 n) \sum_{i \in J} q_i^2 - 2 \lambda_1 \lambda_2 n \sum_{i \in I \cup J} p_i q_i \geq 0. \]

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:
\[ \beta - \text{Div} = \lambda_1 \lambda_2 n \sum_{i \in I} p_i^2 + \sum_{i \in J} q_i^2 - 2 \lambda_1 \lambda_2 n \sum_{i \in I \cup J} p_i q_i = \]
\[ = \lambda_1 \lambda_2 n \sum_{i \in I} (p_i - q_i)^2. \]

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

If the two habitats have no species in common, then \( I \cap J = \emptyset \), \( n = n_1 + n_2 \), and the beta-diversity is:
\[ \beta - \text{Div} = (n - \lambda_1 n_1 - \lambda_2 n_2) + \]
\[ + \lambda_1 (n_1 - \lambda_1 n) \sum_{i \in I} p_i^2 + \lambda_2 (n_2 - \lambda_2 n) \sum_{i \in J} q_i^2. \]

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:
\[ \beta - \text{Div} = \]
\[ = \lambda_1 n_1 \left[ 1 - \sum_{i \in I} p_i^2 \right] + \lambda_2 n_2 \left[ 1 - \sum_{i \in J} q_i^2 \right] = \]
\[ = \lambda_1 n_1 \left[ 1 - \frac{1}{n_1} \right] + \lambda_2 n_2 \left[ 1 - \frac{1}{n_2} \right], \]

which tends to \( +\infty \) if \( n_1 \) tends to \( +\infty \) or \( / \) 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 \( h_1, \ldots, h_n \) 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 \( \text{Sim} \).

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 0.9999999; 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:
\[ \exp(H) = \exp(\ln 150000000) = 150000000 \]
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:
\[ RGS = n - 1 = 150000000 - 1 = 14999999 \]
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:
\[ \alpha - \text{Div} = \frac{1}{2} GS(h_1) + \frac{1}{2} GS(h_2) = 0.99999995, \]
\[ \gamma - \text{Div} = GS(\lambda_1, h_1; \lambda_2, h_2) = \]
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:

\[
\text{Sim} = \frac{0.9999999}{0.99999995} = 0.99999995.
\]

Using the Shannon entropy:

\[
\begin{align*}
H(h_1) &= \ln 300000000 = 17.2167, \\
H(h_2) &= \ln 150000000 = 16.5235,
\end{align*}
\]

which shows 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 \( h_1 \) that are not found in \( h_2 \), and both \( h_1 \) and \( h_2 \) have their species uniformly distributed. These features make \( h_2 \) somewhat similar to \( h_1 \). Using \( \text{RGS} \), the similarity index is:

\[
\text{Sim} = \frac{2.25 \times 10^7}{3 \times 10^7} = 0.75.
\]

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

\[
egin{align*}
\text{RGS}(h_1) &= 0, \quad \text{RGS}(h_2) = 0, \quad \alpha - \text{Div} = 0, \\
\gamma - \text{Div} &= \text{RGS}(1/2, h_1; 1/2, h_2) = 2 \left( \frac{1}{2} - \frac{1}{4} \right) = 1, \\
\beta - \text{Div} &= 1, \quad \text{Sim} = \frac{0}{1} = 0.
\end{align*}
\]

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 \( \text{km}^2 \)) 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 bird species 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 objective here is to calculate the numerical values of the alpha, gamma, and beta diversities, using the quadratic objective here is to calculate the numerical values of the alpha, gamma, and beta diversities, using the quadratic

Table 1. Application: The data set.

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

Table 2. Relative frequency and the $RGS$ index.

<table>
<thead>
<tr>
<th>$h_j$</th>
<th>$p_{1,j}$</th>
<th>$p_{2,j}$</th>
<th>$p_{3,j}$</th>
<th>$RGS(h_j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.777778</td>
<td>0.000000</td>
<td>0.222222</td>
<td>0.691358</td>
</tr>
<tr>
<td>$h_2$</td>
<td>0.076923</td>
<td>0.214286</td>
<td>0.714286</td>
<td>0.873887</td>
</tr>
<tr>
<td>$h_3$</td>
<td>0.250000</td>
<td>0.357143</td>
<td>0.428572</td>
<td>1.000000</td>
</tr>
<tr>
<td>$h_4$</td>
<td>0.222222</td>
<td>0.222222</td>
<td>0.357143</td>
<td>1.000000</td>
</tr>
<tr>
<td>$h_5$</td>
<td>0.368421</td>
<td>0.555556</td>
<td>0.081633</td>
<td>0.222222</td>
</tr>
<tr>
<td>$h_6$</td>
<td>0.826087</td>
<td>0.357143</td>
<td>0.691358</td>
<td>1.813776</td>
</tr>
<tr>
<td>$h_7$</td>
<td>0.653061</td>
<td>0.222222</td>
<td>0.125000</td>
<td>0.873887</td>
</tr>
</tbody>
</table>

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 $n - 1 = 3 - 1 = 2.0$.

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

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

$$\gamma - \text{Div} = RGS(\lambda_1, h_1; \ldots; \lambda_7, h_7) = 3 \times \left(1 - \sum_{j=1}^{7} \left(\sum_{i=1}^{7} p_{i,j}\right)^2\right),$$

$$\beta - \text{Div} = (\gamma - \text{Div}) - (\alpha - \text{Div}),$$

where the weights are:

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

The similarity index is:

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

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

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

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:

$$\alpha - \text{Div} = 1.62633, \quad \gamma - \text{Div} = 1.86103, \quad \beta - \text{Div} = 0.234693, \quad \text{Sim} = 0.873887.$$

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

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

we get the following relative elevation weights:

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

$$\lambda_4 = 0.383378, \lambda_5 = 0.113937, \lambda_6 = 0.003022,$$

$$\lambda_7 = 0.297117,$$

for which we obtain:

$$\alpha - \text{Div} = 1.54668, \quad \gamma - \text{Div} = 1.81972, \quad \beta - \text{Div} = 0.273039, \quad \text{Sim} = 0.849955.$$

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

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

we get the following relative area and elevation weights:

$$\lambda_1 = 0.142857, \lambda_2 = 0.142857, \lambda_3 = 0.285714,$$

$$\lambda_4 = 0.142857, \lambda_5 = 0.142857, \lambda_6 = 0.142857,$$

$$\lambda_7 = 0.142857,$$

for which we obtain:

$$\alpha - \text{Div} = 1.68750, \quad \gamma - \text{Div} = 1.85714, \quad \beta - \text{Div} = 0.285714, \quad \text{Sim} = 0.862069.$$

Table 3. Application: The data set.

<table>
<thead>
<tr>
<th>$h_j$</th>
<th>Cont</th>
<th>IndOc</th>
<th>End</th>
<th>Area (km$^2$)</th>
<th>Elevation(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>135</td>
<td>361</td>
</tr>
<tr>
<td>$h_2$</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>119</td>
<td>396</td>
</tr>
<tr>
<td>$h_3$</td>
<td>7</td>
<td>6</td>
<td>15</td>
<td>1865</td>
<td>828</td>
</tr>
<tr>
<td>$h_4$</td>
<td>6</td>
<td>6</td>
<td>15</td>
<td>2512</td>
<td>3069</td>
</tr>
<tr>
<td>$h_5$</td>
<td>7</td>
<td>1</td>
<td>11</td>
<td>258</td>
<td>905</td>
</tr>
<tr>
<td>$h_6$</td>
<td>19</td>
<td>3</td>
<td>1</td>
<td>172</td>
<td>24</td>
</tr>
<tr>
<td>$h_7$</td>
<td>32</td>
<td>4</td>
<td>13</td>
<td>2236</td>
<td>2360</td>
</tr>
</tbody>
</table>
\[ \lambda_j = \frac{1}{7}, \quad (j = 1, \ldots, 7), \]

we obtain the average values:
\[ \alpha - \text{Div} = 1.21876, \quad \gamma - \text{Div} = 1.75528, \]
\[ \beta - \text{Div} = 0.536528, \quad \text{Sim} = 0.694339. \]

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 \textit{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 \cite{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 \textit{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 \textit{RGS} index and its use in the additive partitioning of the species diversity are analyzed. \textit{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 \cite{25}). The application shows that by using the \textit{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 important factors.

REFERENCES