(Review Article)

Segregation concepts and using generalized diversity indices in environmental communities

Gebriel M. Shamia [1, *] and Ala Eddin Alshareef [2]

[1] Department of Statistics, Faculty of Science, University of Benghazi, Libya.
[2] General Director of the Higher Institute of Science and Technology, Benghazi-Libya.

Abstract

The concept of segregation has received less attention in ecology. Species diversity is a concept that includes the number of species (the total number of possible species) in the community, their abundance, and the individuals who are divided among the species, for which biodiversity is generally considered in three dimensions: within communities (α-diversity), between or among communities (β-diversity), and in the total dataset (γ-diversity). Recently, Shamia proposed using an improved generalized diversity index, which includes special cases of MacArthur's and Hill's indices. Such an index is easy to interpret when all species have equal even abundances resulting in an equivalent number to refer to the output in the community. index, when community weights are unequal relative abundances, has been applied to real data for comparing the index's performance under the segregation concept. Such measure for different sites gives meaningful results of multiplicative partitioning.

Keywords: Alpha diversity; Beta diversity; Gamma diversity; Species diversity; Equivalent numbers

1. Introduction

There are numerous ways of measuring species diversity, and there is an important contestation about which diversity indicators are "easy to interpret". Each level of inventory diversity can be calculated using different diversity indices methods; for more details, see Magurran ([16]; Chapters: 2-4). Alpha, Beta, and Gamma diversities are many of the essential descriptive measures of ecology and conservation biology however there are different approaches for their measurement and the criteria for their evaluation. Jost [9], Keylock [11], Lande [12], Lewontin [13], Magurran [16], Taneja [28], Wilson and Shamida [32], and others provide properties of intuitive alpha, beta, and gamma. Tuomisto [29] reviewed that α-measure and β-measure can be generally expressed as Hill numbers.

Diversity indices are intended to measure the variety in a set of data consisting of various types of objects which can be used in many fields of study to assess the diversity of any population in which each member belongs to a unique group, type, or species. The Shannon [23] index and the Simpson [24] index are still widely used in ecology. For instance, diversity indices are used to measure biodiversity in demography to measure the distribution of the population of various demographic groups (regions).

To describe and compare different communities, ecologists recognize three components of species richness: alpha, beta, and gamma diversities. Alpha diversity is local diversity (the number of species within a generally small area) while gamma diversity is the total regional diversity of a large area (the number of species covering a large area). Beta diversity links alpha and gamma diversity, or local and regional diversity (the change in the species richness between communities).

---

* Corresponding author: Gebriel M. Shamia

Michel [18] argues that the association between alpha and beta must be understood at multiple scales, he suggested that, in practice alpha diversity is a saturating function of gamma diversity due to the interactions between species (competitive) in their environment, hence the balance among alpha and beta diversity.

Recently, biodiversity and species diversity have been widely used terms in ecology. The importance of biodiversity indices will increase with the quantity of records that those indicators provide. So-called raw indices may not provide the necessary information to study the protection and stability of biotic communities. In studies and research that aim to follow the impact of environmental changes on the stability and productivity of ecological communities, we need measures that provide sufficient information about these changes in the composition of underlying communities. Heterogeneity measures combine species richness and evenness to describe the distribution of individuals in a habitat.

There are two types of heterogeneity measures. Type-I heterogeneity measures place the most weight on the rare species in the sample, and the Shannon function is an example of these measures. Type-II heterogeneity measures place the most weight on the common species, Simpson's index is an example. Most species are rare, and therefore efforts to measure and compare biodiversity. This paper recommends using Hill diversity as a unifying concept to link different measures of diversities.

Shamia [26] proposed a general formulation of diversity indices encompassing many well-known diversity indices based on Hill numbers (Equivalent or Effective numbers of species), which is the number of equally abundant species necessary to produce the observed value of diversity. So, it will use this general expression relating to Good [7] and satisfy the desired property as a meaningful measure of variety between communities due to Routledge [22]. Such general formulations of diversity indices are generalizations of Hill [8] and MacArthur [15] indices that can be used to measure α-measure and β-measure.

2. Related Diversity Indices

The use of families of heterogeneity indices has a long history in ecology, see Magurran [16] and Peet [19]. The first family was proposed by Good [7], including Shannon's index and Simpson's index, in the form

$$H(a,b) = \sum_i^s \pi_i{}^a\,(-\log \pi_i)^b;\,(a,b)=0,1,2,\dots,\dots\dots\dots\dots\dots\;\boldsymbol{Eq.\,1}$$

where (a,b) were defined for non-negative integers. It gives more weight to rare than to common species.

Diversity indices summarize the total number of species in a population and the degree of evenness of the relative abundances.

Shamia [25] suggested a generalized diversity index, $H(a,b)$, due to Good's index and discussed its optimal bounds of (a,b) within a closed region in the quadrant $0 \le a \le 2$ and $-2.5 \le b \le 2.5$, for ecological applications.

$$H(a,b) = \sum_i^s \pi_i^a(-ln\,\pi_i)^b;\,(a,b)\in\Re.\dots\dots\dots\dots\dots\dots\;\boldsymbol{Eq.\,2}$$

Here π_i is the relative abundance of taxon-i (species-i), such that $0 \le \pi_i \le 1$, for ecological applications, $\sum\pi=1$; S is the total number of species present in the community and (a,b) define the structures of the algorithm for real values.

Within this framework, for $0 \le a \le 0.3267$ the valid region for rare species is given by $0 \le b \le 4a(1-a)$, while for $0.3267 < a \le 1.00$ the valid region for b satisfies $0 \le b \le 0.153426+0.693147a+\frac{1}{2}\sqrt{(0.094159+2.772589a)}$; say region-1 "$\boldsymbol{R_1}$". For $H(a,b)$ it consists of the region $a \ge 1$ and $b \le 0$ for common species; say region-2 "$\boldsymbol{R_2}$". Our study was based on artificial populations representing species abundance patterns from the *Brocken-stick model* due to MacArthur [14] and the *equiprobable model*. Baczkowski *et al.* [2,3] determine the range of values (a,b) for which $H(a,b)$ satisfies two key properties of diversity measure due to Pielou ([19]; pp.7).

Thus diversity indices may be categorized as those most sensitive to changes in the rare species, an example being Shannon's index, $H(1,1)$, and those indices which are most sensitive to changes in the common species, an example being the inverse or complement of Simpson's index, $H(2,0)$.

In the 1970s of the last century, Hill [8] proposed a family of diversity indices that may be interpreted as mean numbers of species, see Jost [9]. Use Hill's numbers which encompass the diversity measures that quantify diversity in units of equivalent numbers of equally abundant. Hill's numbers have two major advantages over diversity indices: i) the interpretation of the measure and ii) the sensitivity towards abundance. This includes the reciprocal of Simpson's index (Hill's N_2) or the exponential form of the Shannon function (Hill's N_1) to describe heterogeneity. We should decide beforehand whether we wish to weight the common species more or the rare species more) in our community analysis.

Jost [9] showed that the diversity in ecology corresponds not to the value of the diversity index itself but to its number equivalent. These numbers can be written in Hill's form as

$$N_a = \left(\sum_i^s \pi_i^a\right)^{1/(1-a)}; a \geq 0, a \neq 1. \dots \dots \dots \dots \dots Eq.3$$

The parameter "a" determines the sensitivity of Hill's measure to the relative frequencies. When $a = 0$, the abundances of individual species do not contribute to the sum in the above equation. Rather, only presences are counted, so that N_0 is simply species richness. For $a = 1$, it is undefined, but its limit as "a" tends to one is the exponential of the familiar Shannon index:

$$N_1 = \lim_{a \to 1} N_a = Exp\left(-\sum_i^s \pi_i(ln\, \pi_i)\right). \dots \dots \dots \dots \dots Eq.4$$

MacArthur [15] has presented an index of heterogeneity "the Shannon-Wiener index" which may be expressed in units of numbers of species as a Hill's number (N_1).

When $a=2$, N_2 yields the inverse of the Simpson diversity as follows:

$$N_2 = \left(1 \Big/ \sum_i^s \pi_i^2\right). \dots \dots \dots \dots \dots Eq.5$$

This places more weight on the frequencies of abundant species and discounts rare species. Researchers using Hill numbers should at least report the diversity of all species (a=0), "typical" species (a=1), and dominant species (a=2).

Jost [10] modified some row diversity indices in the concept of equivalent numbers:

Shannon entropy, $H \equiv -\sum_{i=1}^s \pi_i\, (ln\, \pi_i)$, numbers equivalent is Exp(H),

Simpson concentration, $H \equiv \sum_{i=1}^s \pi_i^2$, numbers equivalent is $1/H$,

and

Gini-Simpson index, $H \equiv 1 - \sum_{i=1}^s \pi_i^2$, numbers equivalent is $1/(1-H)$.

In this study, we are dealing with an *Improved Generalized Diversity Index* (*IGDI*); $N(a,b)$, which satisfies the key properties due to Pielou ([20], pp.7) and another desirable property due to Routledge ([22], pp.511); high index scores indicate both high species richness and equal distribution of individuals among species.

2.1. Partitioning diversity

The utility of partition is the statistical significance of level-specific alpha and beta estimates. Partitioning total biodiversity into components within and between communities provides a framework for measuring diversity at different levels of organization (Godfray and Lawton [6]; Lande [12]). The most frequently used diversity indices (*i.e.* Shannon's and Simpson's) are based on information theory (Magurran [16], pp.100-130).

Whittaker [31] asserted multiplicative partitioning that gamma equals the product of alpha and beta (and hence beta can be calculated by dividing gamma by alpha) and quantified Alpha (α), Beta (β), and Gamma (γ) diversities to describe diversity at different hierarchical scales, *i.e.*, within plots and between plots, independently. The idea of a number equivalent originated in economics and was first introduced to ecologists by Hill [8].

Partitioning the "true" gamma diversity multiplicatively (rather than additively) into alpha and beta components allows for a unified treatment of not only alpha and gamma but also beta diversity as counts of an effective number of entity types (Routledge, [21, 22]).

Jost [10] points out that "the statistical independence between alpha and beta is primarily an empirical question; it depends on nature and our sampling scheme". Baselga [4] points out that "the empirical tests demonstrated that multiplicative partition of true diversity yields independent alpha and beta components; additive partition does not". Veech and Crist [30] dealt primarily with species richness, ecologists partitioning diversity generally use multiple diversity measures and compare the beta values among them. This partitioning is special for any given measure.

Jost [9] showed that it gives the equivalent numbers of all standard diversity indices and proposed that diversity values be converted into equivalent or effective numbers of species (also known as Hill numbers), that the number of equally abundant species necessary to produce the observed value of diversity. If Dv_α and Dv_β are to be independent, the numbers equivalents of the alpha, beta, and gamma components of a diversity index in form based on Hill numbers must be related by

$$Dv_\gamma = Dv_\alpha \text{ x } Dv_\beta.$$

From there, β-diversity (Dv_β) is obtained by Dv_γ/Dv_α.

3. Theoretical Background

To obtain easily interpretable measures, suppose that the relative abundances of species in the community of S species are $\pi_1, \pi_2,..., \pi_s$. Then the improved generalized diversity index (IGDI) with suitable (a,b) values, which gives more weight to rare than to the common species for ($0 \leq a \leq 1, b \geq 0$). Otherwise, *IGDI* gives more weight to the common than the rare for ($a \geq 1, b \leq 0$).

$$N(a,b) = \left\{\frac{H(a,b)}{(\ln s)^b}\right\}^{\frac{1}{1-a}} ; a \geq 0 \ (a \neq 1), b \in \{-2.5, 2.5\}., \dots \dots \dots \boldsymbol{Eq.\, 6}$$

$$N(1,b) = Exp\{H(a,b)\}^{1/b} ; \ a = 1, b \in \{-2.5, 2.5\}. \dots \dots \dots \dots \dots \dots \boldsymbol{Eq.\, 7}$$

These expressions can be represented as effective numbers, also known as equivalent numbers, which are straightforward to understand. This exponential $N(1,b)$ arises naturally without referencing information theory.

The central role of the exponential quantity gives the measure a privileged place as a measure of complexity and diversity in all sciences. Since a suitable transformation of $N(a \geq 1, b \leq 0)$, such as inverse or others, would not be required to satisfy the properties. It does not need to be borrowed from information theory but arises naturally from this formalism of equivalent numbers. For more details see Alshareef [1]. Here, "a" is the parameter that determines the measure's sensitivity to species frequencies and "b" is the parameter that represents the sensibility of the index to rare in "$\boldsymbol{R_1}$" or to common species in "$\boldsymbol{R_2}$".

When abundant data is available, consider a community (an ecological population) partitioned into several local communities (k-groups), let j=1, 2,..., k; n_j individuals are sampled in community-j. Let i=1, 2, ..., S denote the species that compose the community, n_{ij} the number of individuals of species-i sampled in the local community-j, $n_i = \sum_j^k n_{ij}$ is the total number of individuals of species-i, $n = \sum_i^s \sum_j^k n_{ij}$ is the total number of sampled individuals. Within each community-j, the probability π_{ij} for an individual to belong to species-i is estimated by $\hat{\pi}_{ij} = p_{ij} = n_{ij}/n_j$, where individuals are divided into S species in terms of weights of relative abundances of species. The same probability for the community is π_i. The relative abundance of species-i can be estimated by $\hat{\pi}_i = p_i = n_i/n$; $\sum_i^s n_i = n$. Communities may have a weight, w_j, satisfying $\pi_j = \sum_i^s w_j \pi_{ij}$. The commonly used w_j is the statistical weight of community-j, usually, the number of individuals in community-j divided by the total number of individuals in the region, $w_j = (n_j/n)$. When the community weights are all equal (identical), w_j=l/k $\left(\sum_{j=1}^k w_j = 1\right)$, see Routledge ([22]; pp.507) for more details.

Diversity can be partitioned into α-measure (the diversity of categories represented within a sample), γ-measure (the total diversity of a region), and β-measure (the diversity of sample types). The alpha-diversity, gamma-diversity, and beta-diversity are noted by $N_\alpha(a,b)$, $N_\gamma(a,b)$, and then $N_\beta(a,b)$. Alpha never exceeds gamma.

There are two distinct cases, which are treated separately:

3.1. Case-I: Alpha and beta when community weights may be unequal:

For $a{\neq}1$, Alpha-diversity (*within-group components*) is calculated as

$$N_\alpha(a,b) = \left\{\frac{H_\alpha(a,b)}{(\ln s)^b}\right\}^{\frac{1}{1-a}} ; a \geq 0, \dots\dots\dots\dots\dots\dots \boldsymbol{Eqn.\,8a}$$

where

$$H_\alpha(a,b) = \sum_j^k w_j \left\{\sum_i^s \pi_{ij}^a \left(- ln\,\pi_{ij}\right)^b\right\}.$$

Gamma-diversity (*Total diversity in the dataset*) is calculated as:

$$N_\gamma(a,b) = \left\{\frac{H_\gamma(a,b)}{(\ln s)^b}\right\}^{\frac{1}{1-a}} ; a \geq 0, \dots\dots\dots\dots\dots\dots \boldsymbol{Eq.\,8b}$$

where

$$H_\gamma(a,b) = \sum_i^s [w_1\pi_{i1} + w_2\pi_{i2} + \cdots]^a . [-ln\,(w_1\pi_{i1} + w_2\pi_{i2} + \cdots]^b.$$

So that Beta–diversity (*between-group components*) is given by

$$N_\beta(a,b) = \frac{N_\gamma(a,b)}{N_\alpha(a,b)}. \dots\dots\dots\dots \boldsymbol{Eq.\,8c}$$

For $a=2$ and $b=0$ (inverse Simpson index), we have

$$N_\alpha(2,0) = \{H_\alpha(2,0)\}^{-1} = \left(\sum_j^k w_j \sum_i^s \pi_{ij}^2\right)^{-1} \equiv N_2\ (Hill\ number).$$

For the total Gamma-diversity ($a=2, b=0$), we have

$$N_\gamma(2,0) = \{H_\gamma(2,0)\}^{-1} = \left\{\sum_i^s [w_1\pi_{i1} + w_1\pi_{i1} + \cdots]^2\right\}^{-1} = \frac{1}{\sum_i^s[w_1\pi_{i1} + w_1\pi_{i1} + \cdots]^2}.$$

Then

$$N_\beta(2,0) = \frac{N_\gamma(2,0)}{N_\alpha(2,0)}.$$

This Simpson measure is a meaningful diversity measure which is always positive and not limited to equally weighted communities.

For $a=1$, Alpha-diversity is calculated as

$$N_\alpha(1,b) = Exp\left\{w_1 \sum_i^s \pi_{i1}(-ln\ \pi_{i1})^b + w_2 \sum_i^s \pi_{i2}(-ln\ \pi_{i2})^b + \cdots\right\}^{1/b}, \ldots\ldots\ldots\ldots\ldots\ldots \ \boldsymbol{Eq.\,9a}$$

Gamma-diversity is calculated as:

$$N_\gamma(1,b) = Exp\left\{\sum_i^s [w_1\pi_{i1} + w_1\pi_{i1} + \cdots][-\ln(w_1\pi_{i1} + w_2\pi_{i2} + \cdots]^b\right\}^{1/b} \ldots\ldots\ldots\ldots\ldots\ldots \ \boldsymbol{Eq.\,b}$$

So that Beta –diversity is given by

$$N_\beta(1,b) = \frac{N_\gamma(1,b)}{N_\alpha(1,b)} . \ldots\ldots\ldots\ldots\ldots \ \boldsymbol{Eq.\,c}$$

Equivalent numbers correct the abnormal behaviour of the "similarity or homogeneity" measure described above.

For exponential Shannon index:

$$N_\alpha(1,1) = Exp\left\{w_1 \sum_i^s \pi_{i1}(-l\,n\,\pi_{i1}) + w_2 \sum_i^s \pi_{i2}(-l\,n\,\pi_{i2}) + \cdots\right\}$$

as the Hill's form of Shannon index, and

$$N_\gamma(1,1) = Exp\left\{\sum_i^s [w_1\pi_{i1} + w_1\pi_{i1} + \cdots][-\ln(w_1\pi_{i1} + w_2\pi_{i2} + \cdots]\right\}.$$

Then

$$N_\beta(1,1) = \frac{N_\gamma(1,1)}{N_\alpha(1,1)}.$$

The equivalent numbers of beta for the exponential Shannon index can be transformed into MacArthur's [15] homogeneity measure:

$$N_\beta^{-1}(1,1) = \frac{N_\alpha(1,1)}{N_\gamma(1,1)} . \ \ldots\ldots\ldots.. \ \boldsymbol{Eq.\,10}$$

This measure is a meaningful diversity that satisfies the requisite properties when the community has unequal weights. In addition, this ratio of equivalent numbers corrects the abnormal behaviour of the "similarity or homogeneity" measure.

The relation between the Shannon alpha, beta, and gamma indices follows directly from information theory:

$$H_\alpha(1,1) + H_\beta(1,1) = H_\gamma(1,1).$$

For true Shannon diversity, it can be written that:

$$Exp\{H_\gamma(1,1)\} = Exp\{H_\alpha(1,1) + H_\beta(1,1)\} = Exp\{H_\alpha(1,1)\}.\,Exp\{H_\beta(1,1)\}.$$

Thus necessarily follows Whittaker's [31] multiplicative law; see Jost [9]. Properly, local diversity $N_\alpha(1,1)$ is a special case of *IGDI*, $N(1,b)$.

So that

$$N_\gamma(1,1) = N_\alpha(1,1) \times N_\beta(1,1) .$$

Again for $b \neq 1$, it can be written that

$$N_\gamma(1,b) = N_\alpha(1,b) \times N_\beta(1,b).$$

Based on the given information, it has been demonstrated by the diversity indices above that multiplicative partitioning is suitable in certain aspects. This is true only for $N(a,b)$. The observed α-diversity and β-diversity were computed at each focal scale, where diversity was measured as species richness ($N(0,0)$), exponential of Shannon index, $N(1,1)$ or inverse of Simpson index, $N(2,0)$, to allow comparison with richness, where $N(1,1)$ approximates the number of abundant species and $N(2,0)$ approximates the number of very abundant species.

3.2. Case-II: Alpha and beta when community weights are equally abundant :

For $a \neq 1$, when the k-community with weights w_j are all equal, $w_j = 1/k$ and Alpha component for any diversity index; $Eq(8a)$, $Eq(8b)$, can be written as follows:

$$N_\alpha(a,b) = \left(\frac{1}{(\ln s)^b}\right)^{\frac{1}{1-a}} \left\{\frac{1}{k}\sum_j^k \sum_i^s \pi_{ij}^a (-\ln \pi_{ij})^b\right\}^{\frac{1}{1-a}} ; a \geq 0.$$

or

$$N_\alpha(a,b) = \left(\frac{1}{(\ln s)^b}\right)^{\frac{1}{1-a}} \left\{\frac{1}{k}\sum_j^k H_j(a,b)\right\}^{\frac{1}{1-a}} \ldots\ldots\ldots\ldots\ldots Eq.\, 11a$$

Gamma-diversity (Total diversity in the dataset) is calculated as:

$$N_\gamma(a,b) = \left(\frac{1}{(\ln s)^b}\right)^{\frac{1}{1-a}} \left\{\sum_i^s \pi_i^a(-\ln \pi_i)^b\right\}^{\frac{1}{1-a}} \ldots\ldots\ldots\ldots\ldots Eq.\, 11b$$

Therefore, beta-diversity is given by:

$$N_\beta(a,b) = \frac{N_\gamma(a,b)}{N_\alpha(a,b)}.$$

For $a=1$, it can be written as:

$$N_\alpha(1,b) = Exp\left\{\frac{1}{k}\sum_j^k \sum_i^s \pi_{ij}(-\ln \pi_{ij})^b\right\}^{1/b}, \ldots\ldots\ldots\ldots\ldots Eq.\, 12a$$

or

$$N_\alpha(1,b) = Exp\left\{\frac{1}{k}\sum_j^k H_j(1,b)\right\}^{1/b}.$$

In general, level-specific $N_\alpha(a,b)$ is computed from the equations: $Eq(8a)$, $Eq(9a)$, $Eq(11a)$, and $Eq(12a)$ by aggregating the diversity samples at the level under focus, $N_\gamma(a,b)$ is computed from the equations: $Eq(8b)$, $Eq(9b)$, $Eq(11b)$, and $Eq(12b)$, and then the level specific $N_\beta(a,b)$ are deduced accordingly.

High index scores indicate both high species richness and a more equal distribution of individuals among species. Such transformations express the data in terms of several species and thus are more easily interpreted. To examine the sensitivity of an index to changes in the abundance of species we compare the change in the index value. This is because a reliable index should reflect a significant change in value for even a minor alteration in species abundance. This feature

is crucial for effectively comparing populations with varying species abundance. This can be done by assessing the extent to which the improved generalized diversity index $N(a,b)$ is sensitive to evenness and richness components. By changing the selection of "a" and "b" it is possible to highlight the change in diversity, whether one of richness or degree of evenness.

Now, we will use $N(1,1)$ and $N(2,0)$ to justify the performance of $N(a,b)$ against the new expressions derived by Jost [9] to show the index sensitivity for communities overlap and similarity under the concept of heterogeneity indices.

In this context, we will introduce a more comprehensive measure (IGDI), $N(a,b)$, proposed by Shamia [26]. This measure has useful additional statistical properties without any mathematical transformation (inverse or complementary) regardless of the level of heterogeneity (dissimilarity) measure in the underlying population. In addition, the derivation and acceptance areas for the important properties of the generalized diversity index $H(a,b)$ were achieved, as has been shown in Eq(6) and Eq(7).

These mathematical frameworks of species diversity indices encompassed a lot of well-known diversity indices in ecology. A key aspect of diversity studies is the partitioning into α-diversity and β-diversity (Whittaker; [31]), which explicitly separates measures of biodiversity levels (α-diversity) from measures of biodiversity change (β-diversity); into independent within-group and between-group components.

4. Application

Data were collected from the Benghazi Cancer Registry, in 2003; 997 early-stage cancer cases were registered among residents. Global population-based estimates for all regions combined (excluding non-melanoma skin) are 118 per 100,000 for men and 95 per 100,000 for women; see El-Mistiri et al. [5]. This application checks the segregation property of IGDI by using cancer epidemic data in five departmental divisions: Benghazi, East Benghazi-A, East Benghazi-B, East Benghazi-C, and Southwest Benghazi-D; see Shamia *et al.* [27]. Registration was carried out by active data collection; the registry staff routinely visited all hospitals and pathological laboratories in eastern Libya (1.6 million inhabitants) and collected information from all death registration offices. Here, the types of cancer epidemic represent twenty-three species $S_1, S_2, \ldots, S_{23}$ and the five sites represent the regions. Consider the types of cancer refer to species and the departments represent five regions, where "n_j" is the number of each species in region-j, j=1,2,…, k such that "k" is the number of quadrates under study; where n_1=21, n_2=20, n_3=20, n_4=19, n_5=17, and k =5.

The species relative abundance bars and the pattern broken-stick (Bs) model in the combined region's datasets are shown below.
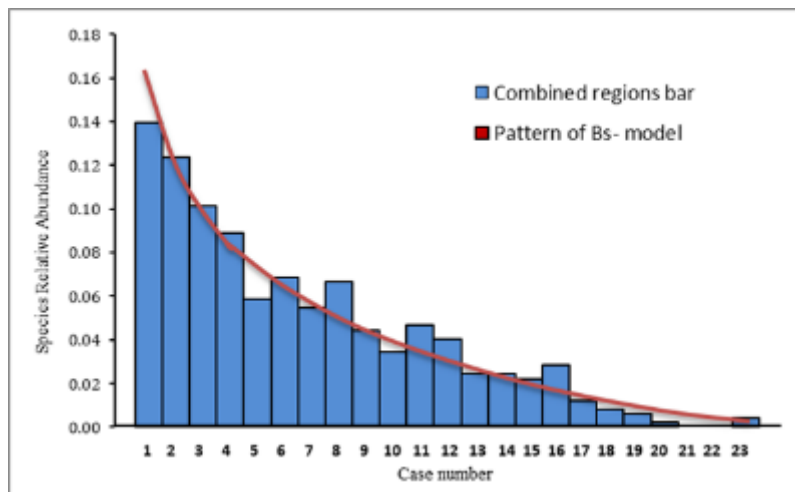

Figure 1 The pattern of the broken-stick model (Bs) and the relative abundances in combined regions.

The results of segregation for this data by applying *IGDI* at different choices of (a,b) within the acceptable regions "R_1" and "R_2" are illustrated in **Table 1**.

Table 1 Gamma, Alpha, and Beta measures at different choices of (a,b)

a	b	γ	α	β	β^{-1}
1.0	1.0	16.301	15.397	1.059	0.945
1.0	1.5	21.498	20.540	1.037	0.965
0.5	1.0	17.858	15.500	1.152	0.868
2.0	0.0	14.032	13.138	1.068	0.936
1.5	0.0	14.111	14.101	1.001	0.999
1.0	-1.0	13.423	12.375	1.085	0.922

From the above table it is clear that total Gamma diversity in the dataset is greater than Alpha diversity within the dataset which means the index satisfies the segregation property, so Beta diversity for the change in species composition between 5 regions represents the maximum value at (0.5,1.0) in the rare area (R_1) and minimum value at (1.5,0.0) in the common area (R_2). For Alpha diversity within the dataset, the maximum value at (1.0,1.5) in the rare area (R_1) and the minimum value at (1.0,-1.0) in the common region (R_2). Again this is true for the total Gamma diversity.

It is clear that for measuring segregation with multiple categories using Exponential Shannon index $N(1.0,1.0)$, we get (N_α =15.397) and (N_β =1.059) and Inverse Simpson index $N(2.0,0.0)$, we have (N_α =13.138) and (N_β =1.068). So the beta diversity for Shannon and Simpson under the concept of equivalent numbers are the beta diversities *Exp* (0.057325) and *Inverse* (0.936283) between 5 regions or communities respectively. It is evident from the above table that the beta diversity estimates for options (a, b) show a significant degree of similarity. This refers to changes in the cancer rates in the dataset of five similar regions.

The homogeneity measure (β^{-1}) for different choices (a,b) in *IGDI* is close to unity. This indicates that all five communities in the dataset are identical (similar) in cancer species composition.

According to the results above the γ-diversities are close to (nearly equal) to the local α-diversities, in general, for our dataset in cancer species composition. So that the cancer species intermingle (intermix) at random. The ratio of the equivalent numbers corrects the abnormal behaviour of the "similarity or homogeneity" measure.

5. Conclusion

Partitioning the total true gamma diversity into the mathematically independent alpha and beta components allows us to check out dataset interiors within and between group diversities. Various measures of species diversity among communities have been proposed, especially for patterns of species richness along environmental gradients.

Very common and more popular indices are MacArthur's measure, $N(1,1)$ and Hill's measure, $N(2,0)$, even if they achieve the agreed basic properties of a good scale.

It appears that there is still limited information about the construction of individuals in the community, in the form of formulas due to Hill [8] and MacArthur [15] in the rare and the common regions (R_1, R_2), that need to be studied. In this context, some mathematical transformations must be applied to make them more sensitive to describing the molecular data.

The paper introduced a method for assessing diversity that considers variety, balance, and disparity across different types. We hope this meets the researchers' requirements. The framework builds on *equivalent numbers* and the corresponding decomposition of the *Improved Generalized Diversity Index*, $N(a,b)$ into independent α and β components.

The results of this study can provide some reference for subsequent studies on the structure and diversity of a group of regions.

We hope that our article and related examples will encourage ecologists to take advantage of Hill's and MacArthur's numbers, which may assist in improving the precision and usefulness of their research.

Compliance with ethical standards

Disclosure of conflict of interest

There are no known competing financial interests or personal interests that could have appeared to influence this paper.

References

[1]   Alshareef, A.E. "A Case Study in the Concept of Biodiversity Measurement: Statistical Properties of an Improved Diversity Index". Unpublished MSc thesis, Department of Statistics, Academy of Graduate Studies, Benghazi-Libya; 2011.

[2]   Baczkowski, A.J., Joanes, D.N., and Shamia, G.M. Properties of a Generalized Diversity Index. Journal of Theoretical Biology.1997; 188: 207-213.

[3]   Baczkowski, A.J., Joanes, D.N., and Shamia, G.M. Range of Validity of (a,b) for a Generalized Diversity Index. Mathematical Biosciences. 1998; 148: 115-128.

[4]   Baselga, A. Multiplicative Partition of True Diversity Yields Independent Alpha and Beta Components; additive partition does not. Ecology. 2010; 91: 1974-1981.

[5]   El-Mistiri, M., Verdecchia, A., Rashid, I., El Sahli, N., El Mangush, M., and Federico, M. Cancer Incidence in Eastern Libya: the first report from the Benghazi Cancer Registry, 2003. Wiley-Liss. 2007; 120(2):392-7.

[6]   Godfray, C. and Lawton, J. Scale and Species Numbers. Environmental Science; TRENDS in Ecology and Evolution. 2001; 16(6): 400-404.

[7]   Good, I.J. The Population Frequencies of Species and the Estimation of Population Parameters. Biometrika. 1953; 40: 237-264.

[8]   Hill, M. Diversity and Evenness: A unifying notation and its consequences. Ecology. 1973; 54: 427-432.

[9]   Jost, L. Entropy and Diversity. Oikos. 2006;113: 363–375.

[10]   Jost, L. (2010): The Relation between Evenness and Diversity. Diversity, 2(2): 2427–2439.

[11]   Keylock, C. (2005): Simpson Diversity and the Shannon-Wiener Index as Special Cases of a Generalized Entropy. Oikos, 109: 203-207.

[12]   Lande, R. Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos. 1996; 76: 5-13.

[13]   Lewontin, R. The Apportionment of Human Diversity. Evolutionary Biology.1972; 6: 381-398.

[14]   MacArthur, R. H.. On the Relative Abundance of Bird Species. Proceedings of the National Academy of Sciences-USA. 1957; 43(3): 293-295.

[15]   MacArthur, R. H. Patterns of Species Diversity. Biological Reviews. 1965; 40: 510 -533.

[16]   Magurran, A. Measuring biological diversity. Blackwell Publishing; Chapter-4, p:100-130. 2004.

[17]   McIntosh, R.P. (1967): The Continuum Concept of Vegetation. Botanical Reviews, 33: 130-87.

[18]   Michel Loreau1 Are Communities Saturated?. On the relationship between α, β and γ diversity. Ecology Letters,. 2004; 3:73-67.

[19]   Peet, R.K. The measurement of Species Diversity. Annual Review of Ecology and Systematics. 1974; 5: 285-307.

[20]   Pielou, E.C. Ecological diversity. John Wiley and Sons, New York, NY, p.7. 1975

[21]   Routledge, R.D. On Whittaker's Components of Diversity. Ecology. 1977; 58: 1120-1127.

[22]   Routledge, R.D. Diversity Indices: which ones are admissible. Journal of Theoretical Biology. 1979; 76: 503-515.

[23] Shannon, C.E. The mathematical theory of communication. En C.E. Shannon; W. Weaver (Ed.): The mathematical theory of communication. Urbana, University of Illinois Press, p: 3-91. 1948

[24] Simpson, E. H. Measurement of Diversity. Nature. 1949;163: 688.

[25] Shamia, G.M. "Statistical Properties of the Generalized Diversity Index". Unpublished PhD. Thesis, Department of Statistics, School of Mathematics, Leeds University, UK; 1997.

[26] Shamia, G.M. Statistical Properties and Performance of Ecological Indices Based on Relative Abundances. Journal of World Academy of Science, Engineering and Technology. 2013; 7(3): 396-401.

[27] Shamia, G.M., El-Mistiri, M., and Bosheha, M. Diversity Indices in Combining Classifiers for Cancer Diseases Data. Journal of Applied Environmental and Biological Sciences. 2012; 2(10): 500-507.

[28] Taneja, I. On Generalized Information Measures and their Applications. Advances in Electronics and Electron Physics. 1989; 76: 327-413.

[29] Tuomisto, H. A Diversity of Beta Diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography, 2010; 33: 23- 45.

[30] Veech, J. and Crist, T.O. Diversity partitioning without statistical independence of alpha and beta. Environmental Science; Ecology. 2010; 91(7), 1964–19691.

[31] Whittaker, R. Evolution and Measurement of Species Diversity. Taxon. 1972; 21: 213- 251.

[32] Wilson, M.V. and Shamida, A. Measuring Beta Diversity with Presence-Absence Data. Journal of Ecology. 1984; 72(3): 1055-1064.