Estimating and comparing biodiversity with a single universal metric

Roberto Cazzolla Gatti\textsuperscript{a,b,*}, Nicola Amoroso\textsuperscript{c}, Alfonso Monaco\textsuperscript{d}

\textsuperscript{a} Biological Institute, Tomsk State University, Tomsk, Russia
\textsuperscript{b} Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria
\textsuperscript{c} Dipartimento di Farmacia - Scienze del Farmaco, Università degli Studi di Bari, Italy
\textsuperscript{d} Istituto Nazionale di Fisica Nucleare – Sez. di Bari, Bari, Italy

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Biodiversity index
Hill numbers
Effective diversity
Absolute diversity

\textbf{ABSTRACT}

The development of a unique index to represent the real diversity of a community is still a challenge. The difficulties to find a single, reliable, diversity metric is mostly due to the fact that the indices allowing an estimation of the number of species actually present in a sample do not provide information on how species abundances are distributed within a community. Although some diversity measures attempted to weigh richness in relation to the total abundance, the homogeneity (evenness) of individuals within species is not integrated into them. The most common diversity indices have been "summarized" by Hill in 1973 in a single equation and, subsequently, different indicators have been proposed to estimate the "absolute" and the "effective" number of species present in a study area in an attempt to find a unified treatment of all standard diversity indices and to compare the true diversity among community. However, the choice of one of the effective number of species indices depends on what aspects of diversity of the study sites the research aims to investigate and could bias the study and mislead the interpretation of the comparative results. Moreover, effective Hill numbers account for only the observed, and not the expected (i.e. the absolute) diversity of the studied community. To compare and interpret the effective number of species with one single metric and analyze the diversity data with techniques that do not depend on a traditional non-parametric index, we developed an absolute measure of diversity based on the effective number of species and derived from the Hill numbers of order 0, 1, and 2. We tested the new index against previously-proposed absolute diversity estimators and the first three Hill numbers taken alone. We simulated the new index's behavior with different gradients of richness, abundance and evenness and we, finally, empirically tested it on tree communities of three biomes (boreal, temperate and tropical forests) of the United States and on a coral reef community of Cuba. This new index proved to be the first reliable and comparable measure, which combines both the absolute richness and the evenness of a community related to the most used traditional indices in a unique, simple and comprehensive numerical value that would represent the absolute effective diversity (AED) of any biological community.

\section{1. Introduction}

The identification of a single ideal biological diversity index, which can be applied to every study and ecosystem, is one of the main aims of ecologists (Magurran, 2013; Cazzolla Gatti, 2017). The simplest metric, species richness, although providing an exact number of species in a sample, can be scarcely informative and imprecise (Magurran & McGill, 2011), being subjected to the random changes in species composition more than any other index (Gotelli & Colwell, 2011). Abundance-based diversity indices add the information of the number of individuals to that of species richness but they are unable to represent the true "diversity" of a community (Chao et al., 2006).

The difficulties to find a unique and reliable diversity metric are mostly due to the fact that richness indices allow a bias-reduced estimation of the number of species actually present in a sample but do not provide any information on how individuals are distributed within species (Béguinot, 2017). Although some measures (e.g. Margalef and Menhinick indexes) attempt to weigh richness in relation to the total abundance, the homogeneity (evenness) of individuals within species is not integrated into them (Gamito, 2010; Chen & Zhang, 1999). The metrics that incorporate a measure of heterogeneity, together with one of richness, are called "diversity indices" (Heip et al., 1998) and those of Shannon and Simpson (sometimes referred to as the Gini-Simpson) are the most used (Simpson, 1949; Spellerberg & Fedor, 2003).

Nevertheless, they cannot be considered indices of “true diversity” because both are highly nonlinear and their values are not comparable...
in terms of the factual number of species and their abundance in a community (Tuomisto, 2010).

Both Shannon and Simpson’s diversity indices are special cases of entropy measures. These indices have been "summarized" by Renyi (1961) and then “formalized” by Hill (Hill, 1973) in a single equation that can help to get their precise description by simply changing the parameter α:

\[
H_\alpha = \left[ \sum_{i=1}^{S} p_i^\alpha \right]^{1/\alpha}
\]

where \(p_i\) is the proportion of individuals belonging to the ith species, \(S\) is the total number of species, and \(\alpha\) an integer value ranging from \(\pm \infty\). If we replace 0 to \(\alpha\) in this equation we obtain \(H_0 = S\), which is the simple number of species \((S)\); \(H_1\) is not defined for \(\alpha = 1\), nevertheless, the limit exists and, when \(\alpha\) tends to 1, \(H_1 = \exp H'\), which is the exponential form of the Shannon index \((H)\); for \(\alpha = 2\), we obtain \(H_2 = 1/D\), that is the reciprocal of the Simpson index \((D)\). Finally, for \(\alpha\) tending to \(\pm \infty\), we get the reciprocal of dominance/rarity indices. In practice, low values of \(\alpha\) tend to give higher weight to rare species, while high values give more weight to dominant ones (Jost, 2006, Cazzolla Gatti, 2018).

It follows that Shannon diversity of order \(\alpha = 1\) \((H_1)\) gives more importance to less abundant species, while Simpson diversity \(\alpha = 2\) \((H_2)\) to dominant ones. When \(\alpha = 0\) \((H_0)\), diversity corresponds to the whole species richness \((S)\) of the sample, because rare species are weighted as well as abundant ones (Jost, 2007). The high correlation between these diversity measures should not be surprising as they can be derived from the same basic generalization entropy formula, being aspects of the same ecological pattern (Chao & Jost, 2015).

Moreover, a novel group of indices has been developed to take into consideration the rare species, which are less likely to be detected and included in the sampling (Chiariucci et al., 2018). Different indicators were also proposed to estimate: (i) the “absolute” and (ii) the “effective” number of species present in a study area. The absolute diversity indices are based on the mark-release-recapture (MRR) statistics-like ratio to estimate the richness by adding a correction factor to the observed number of species (Chao, 1987). Differently, the effective number of species is the number of equally abundant species needed to obtain the same mean proportional species abundance as that observed in the dataset of interest (Jost, 2006).

1.1. The absolute diversity indices

One of the most reliable absolute diversity indices, which considers the rarity of the species represented by one individual (called singletons) and two individuals (doubletons), is the Chao1, developed by Anne Chao (Chao, 1984). The Chao2 is the same index applied to incidence-based data. Chao1 estimates the absolute number of species in a sample, based on the number of rare species, as:

\[
S_{\text{Chao1}} = S_{\text{obs}} + \frac{N_2^2}{2N_1}
\]

where \(S_{\text{obs}}\) is the number of species observed in the sample, \(N_1\) the number of species represented by a single individual (singletons), and \(N_2\) those represented by two individuals (doubletons). In the bias-corrected form, 1 is subtracted from \(N_1\) and 1 is added to \(N_2\).

It is also possible to calculate the variance of this index by the following formula:

\[
\text{var}_{\text{Chao1}} = N_2 \left[ \frac{1}{2} \left( \frac{N_1}{N_2} \right)^3 + \frac{N_1}{N_2} + \frac{1}{4} \left( \frac{N_1}{N_2} \right)^4 \right]
\]

for \(N_1 > 0\) and \(N_2 > 0\).

The main problem with this estimator is that it is strongly influenced by the sample size, so it is important to ensure adequate sampling in order to not underestimate the richness (Colwell & Coddington, 1994; Messina et al., 2016). Moreover, the Chao1 index provides the minimum of richness estimates and assume homogeneity among the samples. This must be remembered, as it would be inappropriate to use it to compare the estimated richness of sites that have a large difference in their specific composition (Magurran, 2013).

1.2. The effective number of species

At the same time, the benefits to derive "effective number of species" from the canonical Shannon and Simpson diversity indices have gained support (Jost 2006). This is because the true diversity of a community can also be considered the number of equally-common species required to give a particular value of an index (Jost, 2006). This approach has an intuitive interpretation. In fact, after the conversion into common units of standard diversity indices, a meaningful comparison of different indices applied to the same area can be performed. This makes possible a unified treatment of all standard diversity indices and to compare the true diversity among community assuming that all species are equally abundant.

The effective number of species can be derived from the following formulas:

\[
H_0 = \sum_{i=1}^{S} p_i^0 = S
\]

\[
H_1 = \exp \left( -\sum_{i=1}^{S} p_i \ln p_i \right) = \exp H'
\]

\[
H_2 = 1/D = \sum_{i=1}^{S} p_i^{-1} = 1/D
\]

where \(S\) is the number of species in the sample, \(p_i = n/N\) is the relative abundance, i.e. the abundance of each species \((n)\) in the sample divided by the total abundance of all species in the sample \((N)\), \(H'\) is Shannon Diversity Index, and \(D\) is Simpson Diversity Index.

These effective number of species, calculated from Hill’s \(\alpha\) values, is a continuum that ranges from the rarest (\(\alpha\) close to 0) to the dominant species (\(\alpha\) towards \(+\infty\)) (Fig. 1).

However, previous attempts to identify an ideal diversity measure have failed to find a unique index to represent the real diversity of a community and, instead, evidenced the need to report at least two measures (Whittaker, 1972; Stirling and Wilsey, 2001; Morris et al., 2014). The multiple effective diversity measures included in Hill’s continuum provides a more comprehensive understanding of how the shift in rare and abundant species drives biological communities (Morris et al., 2014). The main benefit of using these diversity metrics as measures of the true (or absolute) diversity with respect to the closely related and more traditional indices (S, Shannon, and Simpson) is the simplified interpretation of results because units are always in effective number of species regardless of the position along the series (Jost, 2006). Moreover, effective species numbers behave as expected when diversity is doubled or halved, while the traditional indices of diversity do not (Jost, 2006).

Therefore, the choice of one of the effective number of species indices \((H_0, H_1, H_2)\) depends on what aspects of diversity of the study sites the research aims to investigate (Morris et al., 2014). If the effects are expected to be more apparent in rare species or traits, then \(H_0\) would be appropriate. However, if dominant species or traits are expected to be more important, then \(H_2\) would be more relevant. \(H_1\) could be used, instead, in situations where rare and abundant species or traits are expected to be equally important (Morris et al., 2014).

Although converting Shannon and Simpson diversity indices into true diversities (i.e. effective numbers of species) gives them a set of common behaviors and properties (Leinster & Cobbold, 2012), the choice of the most appropriate one could bias the study and mislead the interpretation of the comparative results. Moreover, after conversion, diversity is always measured in units of species number, irrespective
from which index has been used, but it accounts for only the observed, and not the expected (or the absolute; Chao, 1987) diversity of the studied community.

To compare and interpret the effective number of species with one single metric and analyze the diversity data with techniques that do not depend on a traditional nonlinear index, we developed an absolute measure of diversity (Chao, 1984) based on the effective number of species (Jost, 2006) and derived from the Hill numbers (H) of order 0, 1, and 2 (Hill, 1973). This index would constitute the first comparable measure, which combines both the absolute richness and the evenness of a community related to the most used traditional indices (such as S, Shannon, and Simpson) in a unique, simple and comprehensive numerical value that represents the absolute effective diversity (hereafter, AED) of any biological community.

2. The new absolute effective diversity (AED) index

In the form of a Chao-like absolute diversity (i.e. observed species + less abundant/more abundant), our index is composed by the effective number of species (H0) plus a proportion between the exponential of Shannon Index (H1) and the inverse of Simpson Index (H2).

Our new absolute diversity index, built on the effective number of species of a community (hereafter, AED), takes the following simple equation:

$$AED = H_0 + \frac{H_1^2}{2H_2}$$  \hspace{1cm} (7)

and is constructed in a form that sums up the effective richness (H0) to a ratio of less dominant (H1) and more dominant (H2) effective species. This new index is quite convenient since it encapsulates the information of both the three main Hill Numbers, which well describes the diversity profile of a community, in an absolute diversity measures (i.e. Chao-like index), which has been shown to be reliable and less affected by estimation errors than the three traditional diversity measures, on which it is based, taken alone (Magurran and McGill, 2011; Cazzolla Gatti and Notarnicola, 2018).

As well as for the Chao1 index, the AED variance (standard error, S.E.) can be calculated as:

$$S.E._{\text{AED}} = H_0 \left[ \frac{1}{2} \left( \frac{H_1}{H_2} \right)^2 + \left( \frac{H_1}{H_2} \right)^3 + \frac{1}{4} \left( \frac{H_1}{H_2} \right)^4 \right]^{\frac{1}{2}}$$  \hspace{1cm} (8)

For $H_1 > 0$ and $H_2 > 0$ and $n =$ total number of samples. Our measure encapsulates the observed richness and the estimated “missing” diversity due to rare vs. dominant species in a single metric.

In fact, even though dominant species have a greater effect on observed geographical patterns of diversity than do rare ones (Vázquez and Gaston, 2004), the typically larger numbers of rare species (Gaston, 1994; Magurran & Henderson, 2003), the positive correlations between the numbers of rare species and the overall numbers of species in an area (Gaston and Blackburn, 2000; Cazzolla Gatti et al., 2017), and the evidence of nested species distributions (Wright et al., 1998) emphasized the importance of rare species in biodiversity studies and as one of the main targets for conservation (Purvis and Hector, 2000; Lennon et al., 2004; Gatti et al., 2018; Cazzolla Gatti and Notarnicola, 2018). From a biological perspective (Keylock, 2005), if a study is especially concerned with the dominant species, it is better to use higher-order measures. The higher the order, the more the measure emphasizes the commonest species (as depicted in Fig. 1). When the rarest elements of a sample are as important as the commonest elements the diversity of order zero (H0), species richness, is analogous of the species observed in the sample (Sobs in Chao’s formula [Eq. 2]). Along the continuum range of Hill’s measures (Fig. 1), diversity of order one (H1, the exponential of the Shannon entropy) is the first index that accounts for more rare species (analogous to singletons in Chao’s formula) and diversity of order two (H2, inverse of Simpson) accounts for the second group of rare species (analogous to doubletons in Chao’s formula). Higher-order diversities (H3, H4, …) exist but are seldom used and also not considered in Chao’s estimators and similar formulas (such as Jackknife) because the relationship between the first two order rarities (singletons/H1 and doubletons/H2) already proved to be a reliable measure of “missing species” from a sample. Jost (2006) clearly mentioned that the three indexes of diversity of order zero (species richness), one (exponential of Shannon–Wiener index) and two (inverse of Simpson index) provides more information about the samples than any single measure and it is advisable to give all three in any study of a single community so to judge the degree of dominance in the community by looking at the drops between each one. Besides confirming Jost (2006)’s statement on the fact that showing a continuous range of diversities (Fig. 1) and presents a graph of the results (as done by Hill, 1973) is even a better approach, because this gives a clearer graphical picture of the degree of dominance in the community, we have formulated an index capable to include all the main three diversity measures in a way (a Chao-like index) that allows not only a calculation of the effective diversity (invoked by Jost, 2006) but also of the absolute diversity of a community in a single universal metric.

3. AED ecological properties

The AED Index better accounts for the effective richness in a community than traditional indices (Shannon and Simpson Inverse) and fits more with the shape of the effective richness curve while accounting for the “missing species” positioning, therefore, upper (Fig. 2).
The AED Index keeps the linearity of the effective number of species ($H_0$, $H_1$, $H_2$) while accounting for the “missing species” (Supplementary Fig. 1) and shows a higher slope ($m = 1.5$). AED, which is a more comprehensive measure - compared to the other indices - that merges the information of richness to the information of evenness and also accounts for rare species, seems capable to estimate the absolute number of species in a community with a lower amount of input data (in this case: richness).

We also analyzed the behavior (Supplementary Fig. 2) of the three main effective number of species indices ($H_0$, $H_1$, and $H_2$) and the newly proposed Absolute Effective Diversity (AED) Index when plotted on an incremental evenness (from 0 to 1, calculated as the mean between Shannon and Simpson Evenness indexes), keeping the species richness ($H_0$) constant at 100 (the opposite situation of Fig. 2).

Then, we simulated a system with number of species, overall abundance and species evenness kept constant, but with varying the relative abundance of species (Supplementary Fig. 3).

The AED Index does not show any specific trend and this confirms that it is independent of the relative abundances of species (i.e. there is no bias towards specific rare or dominant species).

Moreover, we assessed statistically the absence of any trend. In fact, we evaluated the significance of a linear model $y = ax + b$ describing the AED behavior and found that $a = 0$ and $b = 100$ (within a 95% confidence interval).

We also tested the case of changes in the overall abundance with number of species kept constant and evenness tending to 1. We simulated a variation of the overall abundance from $S$ to 250 individuals (Supplementary Fig. 4), as with $S$ species the overall abundance should at least be equals to $S$.

This simulation shows that when the overall abundance ($A$) is different among the different communities (or within the same community during the time) with $S$ constant, the AED Index does not show significant variations (i.e. it is not influenced by the sample size).

Finally, we compared the new AED Index to previously-proposed absolute diversity estimators, such as the Chao1, Bootstrap, Jackknife of first and second order, and Michaelis–Menten indices as defined in Magurran and McGill (2011) and Magurran (2013) (Fig. 3).

The AED Index shows a performance in line with the other indices, a reduced variability (lower and upper quartiles closer to the median) and a lower number of outliers (less over- and under-estimations). Moreover, between 60–70% of sample coverage, the AED is already able to estimate the absolute number of species in a community, where other metrics (in particular the Chao 1 and Bootstrapping, which performs better than Jackknife and MM, both according to our simulation and the literature) need a sample coverage of 80–90% to provide reliable estimates.

4. Two empirical tests for the AED index

After simulating the AED Index’s behavior with different conditions of $S$, $A$, and $E$ and in a comparison with other estimators, we tested on empirical data the reliability of the new metric in comparing the estimated richness of sites that have a large difference in their specific composition.

First, we estimate the absolute effective tree diversity of three biomes (boreal, temperate and tropical forests) of the United States from data included in the Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service (https://www.fia.fs.fed.us/). Tree data in the FIA database are collected in plots with a 7.32 m radius and d.b.h. > 12.7 cm (5 inches).

Initially, to check if the new index was able to differentiate the three biomes on the basis of their absolute effective diversity, we calculated the traditional diversity indices, their effective version, and the AED Index from a small comparable sample size (53–55 randomly chosen plots). Results (Table 1 and Supplementary Fig. 5) indicate that AED well discriminates the diversity of the three US forest types and can estimate the expected difference in terms of absolute diversity, which is higher in tropical forests (~60 effective species), intermediate in temperate forests (~15 effective species), and low in boreal forests (~4 effective species).

Secondly, after confirming AED Index’s ability to discriminate communities and differences in their diversities even with a small sample size, we tested its performance vs. those of other common absolute richness metrics (namely, Chao1, Jackknife1 and 2, Bootstrapping and Michaelis–Menten) in each of the three forest types for the whole large FIA datasets (Fig. 4).

This allowed us to check whether the AED Index was reliable in comparison with other estimators and to better understand its properties in estimating the absolute effective diversity of a community.

Accumulation curves show that the AED Index offers a good balance between under- and over-estimations, and the ability to account for the real number of species with a lower sampling effort (especially with respect to Chao1 and Jackknife 1). In particular, in more diverse communities, such as tropical and temperate forests, the AED curves are
placed between those (higher) of Jackknife 2 and Bootstrapping and those (lower) of Chao1 and Jackknife1.

In tropical forests, the AED Index requires an extremely lower sampling effort than other indices.

In tropical forests, the AED Index is able to estimate the real diversity of the community with half of the sampling effort needed for the Chao1 and a similar sampling effort needed for Jackknife 1 and 2 indices, whereas the Bootstrapping requires a lower sampling effort but excessively overestimates the absolute diversity of the community.

Significantly, when the community has a low species richness and the sampling effort required to estimate the real diversity is lower, as for the case of US boreal forests (\(S_{\text{absolute}} = 9\)), the AED gets this value after accumulating just \(\approx 60\) random plots (compared to more than hundreds needed by the other indices).

As an additional test on a different ecosystem, we estimate the absolute and effective coral community diversity of a reef at Naval Station Guantánamo Bay in southeastern Cuba from a published dataset (Chiappone et al., 2001). In this study, 25 m transects were oriented perpendicular to shore on four haphazardly selected spurs and used to randomly select \(1 \times 1\) m quadrat locations. Abundance data in terms of area coverage (%) were collected per each species (including, octocorals, stony corals, turf algae, coralline algae, etc.).

As above, we tested AED performance vs. those of other common absolute richness metrics (namely, Chao1, Jackknife1 and 2, Bootstrapping and Michaelis-Menten) on this coral reef community (Fig. 5).

Again, in this test, the AED Index requires a low sampling effort (40% coverage) to estimate the absolute diversity if the community.

Consistently, the AED Index shows higher accuracy than other indices in the estimation, with a lower variability around the mean and no extremely over- and under-estimations (outliers).

### Table 1

The estimated species diversity (Shannon and Simpson indices) and effective richness (Hill and AED indices) in a small and comparable sample of three forest types in the United States (data from the FIA database).

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Boreal mean (n. 53)</th>
<th>SD</th>
<th>Temperate mean (n. 54)</th>
<th>SD</th>
<th>Tropical mean (n. 55)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon</td>
<td>0.74</td>
<td>0.23</td>
<td>1.37</td>
<td>0.50</td>
<td>2.08</td>
<td>0.60</td>
</tr>
<tr>
<td>H1</td>
<td>2.16</td>
<td>0.50</td>
<td>4.41</td>
<td>2.13</td>
<td>9.23</td>
<td>4.40</td>
</tr>
<tr>
<td>Simpson</td>
<td>0.46</td>
<td>0.13</td>
<td>0.65</td>
<td>0.18</td>
<td>0.80</td>
<td>0.14</td>
</tr>
<tr>
<td>H2</td>
<td>2.63</td>
<td>2.15</td>
<td>1.78</td>
<td>0.95</td>
<td>1.32</td>
<td>0.46</td>
</tr>
<tr>
<td>Richness/H0</td>
<td>2.83</td>
<td>0.85</td>
<td>6.14</td>
<td>2.68</td>
<td>13.13</td>
<td>5.87</td>
</tr>
<tr>
<td>AED</td>
<td>4.08</td>
<td>1.62</td>
<td>15.35</td>
<td>12.34</td>
<td>59.22</td>
<td>46.50</td>
</tr>
</tbody>
</table>

Consistently, the AED Index shows higher accuracy than other indices in the estimation, with a lower variability around the mean and no extremely over- and under-estimations (outliers).

### 5. AED vs. other “effective” metrics

The AED completely controls community structure and avoid biases related to varying sample sizes. However, because it is well-known that relationships between diversity indices do not always follow mathematically predicted patterns (Stirling and Wilsey, 2001; Nagendra, 2002), we also performed some analyses on real data to ensure that our conclusions would be empirically valid.

From our simulations and empirical tests, we observed that AED is always above the H\(_2\) curve and this reflects its purpose to account for additional rare species not included in the observed species of the sampled community. Both H1 and H2 values are below H0 and this is due to the fact that the effective number of species calculated from the Shannon and Simpson diversity indices is always lower than H0 (Fig. 1) because they give more weight to abundant and dominant species. H1 and H2 curves tend to H0 at about 0.6 and 0.9 (Supplementary Fig. 2), respectively, as expected by their placement along with the continuum range shown in Fig. 1.

H1 and H2 always show lower values of “effective number of species” than H0 at low evenness (particularly when E < ≈0.6) and, in this way, they could be highly biased and underestimate the contribution of the “missing effective” species to the true diversity of a community.

The AED curve takes the shape, along an incremental evenness from 0 to 1, of a normal distribution, which would well represent the normal distribution of all species in a community (Connolly et al., 2005; Dornelas et al., 2006; Matthews & Whittaker, 2015). This curve reaches its maximum value at an intermediate evenness level (≈0.5), reasonably because at this evenness both rare and dominant species are almost equally represented and the AED index accounts for a higher number of “missing effective rare” species to add to the effective species richness (H0). Similarly, when the evenness tends to 0, the dominance increases and the number of missing effective rare species would not add much to the...
diversity of the sample calculated as $H_0$.

Finally, comparing the AED index with the Chao1 estimator from which it takes the form, they show similar performances with, at least, three main advantages for AED: its accuracy is much higher (a very low variability around the mean), it shows no extreme over- and under-estimations (no far outliers), and it is able to estimate both the effective and absolute diversity of a community from just the three main diversity indices ($S$, Shannon, and Simpson) without the need to account for singletons and doubletons. In fact, the information on $S$, Shannon and Simpson indices is almost always available in the literature (and can be used to produce estimates from previously published study) and it is always quantifiable from abundance-based datasets, whereas the number of singletons and doubletons may not be present in published data and in actual samples (it is common the case when either singletons or doubletons or both are $= 0$), making the estimation of the absolute species richness from the Chao1 index impossible.
Fig. 5. A comparison among the most common absolute richness estimators (Chao1, Jackknife 1 and 2, Bootstrapping and Michaelis–Menten $S_n$) and the new AED Index in the coral reef community from Chiappone et al., 2001. The continuous horizontal line represents the observed number of species. Estimates were randomized 100 times.

6. Conclusion

The indexes of the effective number of species ($H_0$, $H_1$, and $H_2$) perform well only when the evenness of a sampled community is close to 1 (i.e. all species are almost equally abundant or rare). However, when the evenness is lower than 1 and between 0 and 0.6, which is the case of many real natural communities (see the index performance on real ecosystems, Table 1 and Figs. 4 and 5), the new AED Index is better able to account for the “missing effective” species that contribute to the diversity of the community.

Furthermore, AED represents a reliable estimator of the true diversity of a community (since it is based on Hill numbers, which also take into account the evenness) and not just an estimator of species richness. Finally, we believe that the main strength of this index is that it can be used to simply compare the true diversity of communities, ecosystems, and biomes worldwide with unitless, effective and absolute values.

CRediT authorship contribution statement

Roberto Cazzolla Gatti: Conceptualization, Methodology, Data curation, Validation, Writing - original draft, Writing - review & editing.
Nicola Amoroso: Methodology, Software, Data curation, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.
Alfonso Monaco: Methodology, Software, Data curation, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials


References


