Accepted Manuscript

Using aggregation functions to model human judgements of species diversity

Gleb Beliakov, Simon James, Dale G. Nimmo

PII: S0020-0255(15)00106-1
DOI: http://dx.doi.org/10.1016/j.ins.2015.02.013
Reference: INS 11406

To appear in: Information Sciences

Received Date: 9 October 2014
Revised Date: 1 February 2015
Accepted Date: 6 February 2015

Please cite this article as: G. Beliakov, S. James, D.G. Nimmo, Using aggregation functions to model human judgements of species diversity, Information Sciences (2015), doi: http://dx.doi.org/10.1016/j.ins.2015.02.013

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Using aggregation functions to model human judgements of species diversity

Gleb Beliakov, Simon James and Dale G. Nimmo

Deakin University
School of Information Technology (G. Beliakov and S. James)
School of Life and Environmental Sciences (D. G. Nimmo)
221 Burwood Hwy, Burwood, Victoria 3125, Australia

Abstract

In environmental ecology, diversity indices attempt to capture both the number of species in a community and the relative abundance of each. Many indices have been proposed for quantifying diversity, often based on calculations of dominance, equity and entropy from other research fields. Here we use linear fitting techniques to investigate the use of aggregation functions, both for evaluating the relative biodiversity of different ecological communities, and for understanding human tendencies when making intuitive diversity comparisons. The dataset we use was obtained from an online exercise where individuals were asked to compare hypothetical communities in terms of diversity and importance for conservation.

Keywords: Aggregation functions, weights learning, Bonferroni mean, species diversity, ecological indices,

1Corresponding author. Tel.: +61 3 9251 7481.
E-mail addresses: gleb@deakin.edu.au (G. Beliakov), sjames@deakin.edu.au (S. James), dale.nimmo@deakin.edu.au (D. G. Nimmo).
Using aggregation functions to model human judgements of species diversity

Gleb Beliakov, Simon James\textsuperscript{1} and Dale G. Nimmo

\textit{Deakin University}
\textit{School of Information Technology (G. Beliakov and S. James)}
\textit{School of Life and Environmental Sciences (D. G. Nimmo)}
\textit{221 Burwood Hwy, Burwood, Victoria 3125, Australia}

\textsuperscript{1}Corresponding author. Tel.: +61 3 9251 7481.
E-mail addresses: gleb@deakin.edu.au (G. Beliakov), sjames@deakin.edu.au (S. James), dale.nimmo@deakin.edu.au (D. G. Nimmo).

\textit{Preprint submitted to Elsevier February 10, 2015}

1. Introduction

When policy makers and governments talk about the importance of preserving biodiversity, they have an intuitive rather than mathematical idea of what diversity is. Models that inform management decisions and ecology research, however, require notions like species diversity to be quantified. The debate surrounding which indices should be used is often argued with reference to some key properties and simple examples that show one or another index to be counter-intuitive, however there is no real basis upon which one can be deemed as more acceptable than another, so decision makers are able to choose the index that best serves their agenda. Although communities with no diversity (only one species present) and perfect diversity (a high number of species all equally abundant) can be easily defined, comparisons in diversity toward the middle of the spectrum are more difficult to formally articulate, and to date, there has been no study of whether existing models generally reflect the human perception of what it means for a forest or grassland area to be high or low in biodiversity. In this article, we apply the theory of aggregation functions in order to uncover the implicit judgements that are made when humans evaluate species diversity. We provide a method for making diversity assessments based on samples of elicited comparison judgements from experts, and use a real data set to show that
the resulting evaluations correspond better with intuitive human perceptions than some of the popularly used diversity indices.

Aggregation functions have been successfully employed across a broad range of human decision-making contexts to uncover the influence and importance of variables and predict human judgements, e.g. in journal ranking [2], group decision-making [23, 12] and image analysis [16].

The real data used was gathered from an online exercise we conducted where participants were asked to intuitively judge whether one community was more abundant than another. To inform their decision, they were provided only with the information that is currently used to evaluate species diversity, namely, the individual abundances for all species present in the community\(^2\). From these supplied pairwise comparisons, we can see how well the indices used in ecology literature reflect these judgements, and further, whether results from the theory of aggregation functions and weight learning could potentially provide biodiversity evaluations that are closer to our human perceptions. Many of the evaluations for evenness and diversity that have been developed and used in ecology can themselves be considered in the framework of aggregation functions.

We found that from the simple comparison data obtained from our online survey, aggregation functions could be constructed that were more consistent with the human evaluations than any one index alone. These functions aggregated the following indices:

- **total abundance** - the number of individuals present or observed across all species in the community;
- **richness** - the number of different species-types;
- **evenness** - an index reflecting how evenly the different species are distributed, which is usually calculated from richness and the individual abundances;
- **species diversity** - an index that usually is said to incorporate both richness and evenness, commonly calculated as their product.

\(^2\)We make the special note that our perception of diversity may be dependent on other factors, such as which species, how rare they are, how different they are to one another, and so on, however none of the current indices take these factors into account.
Even when we only include total abundance, richness and evenness in the aggregation, we can still produce better results than the existing species diversity indices in terms of modeling the human assessments. In particular, we found that weighted Bonferroni means had the best overall performance, which can be explained by their ability to model mandatory requirements, i.e. they can model semantics such as:

IF richness is high and the average of evenness and abundance is high, THEN the community is high in biodiversity.

The advantage of using parameter-learning over an axiomatic approach is that we can model the subtle intuitive trade-offs that are made by experts without having to specify them a priori. In quantifying biodiversity, there is clearly a trade-off that occurs between having a high number of species present with some in low abundance, and having a lower number of species whose abundances are all equal. Although the aggregation functions used can still be interpreted in terms of reasonable properties and behavior, determining their parameters according to expert opinions allows a further justification in that they conform to human expectations and intuitions.

The article will be set out as follows. In Section 2, we present the preliminary concepts required for conducting our analysis. To this end we provide the background on aggregation functions and how to learn their weight parameters using linear programming. In Section 3, we give an overview of ecological indices used for quantifying species diversity and provide details of the online exercise we conducted and the dataset obtained. We then show how aggregation functions can be used to predict human judgements of species diversity in Section 4, and then show the analysis side of this process, i.e. how aggregation functions can be used to interpret the tendencies of the respondents, in Section 5. In Section 6 we provide some discussion and indicate some of our future research directions and finally Section 7 concludes.

2. Preliminaries

We firstly set out the necessary background in aggregation functions which we will use to formulate evaluations of species diversity. After giving some of the important definitions, properties and families, we provide an overview of how we can fit aggregation functions to data along with some recent results and techniques.
2.1. Aggregation functions

Aggregation functions lie at the heart of many decision processes where it is necessary to summarize an input set with a single value. Although many of the results have arisen independently in various research fields, overviews of construction methods and properties with a particular focus on soft computing and decision making can be found in [9, 14, 21].

We will use the following definition.

Definition 1. An aggregation function \( f : [0, 1]^n \to [0, 1] \) is a function non-decreasing in each argument and satisfying \( f(0, \ldots, 0) = 0 \) and \( f(1, \ldots, 1) = 1 \).

Aggregation functions can be defined for inputs other than those given over the unit hypercube, however for the moment we will restrict our considerations to this case. Scaling is employed where necessary in our fitting procedures to ensure that all inputs are given over the same interval.

Depending on the context, the aggregation function desired may belong to a specific class.

Definition 2. With respect to a multivariate input vector \( \mathbf{x} = (x_1, \ldots, x_n) \), an aggregation function \( f \) is considered to be: averaging where for all \( \mathbf{x} \) it holds that \( \min(\mathbf{x}) \leq f(\mathbf{x}) \leq \max(\mathbf{x}) \), conjunctive if \( f(\mathbf{x}) \leq \min(\mathbf{x}) \), disjunctive if \( f(\mathbf{x}) \geq \max(\mathbf{x}) \), and mixed otherwise.

Due to the monotonicity of aggregation functions, averaging behavior is equivalent to idempotency, i.e. \( f(t, t, \ldots, t) = t \).

In particular, we are interested in the following families, all of which are averaging and defined with respect to weights allowing different importances to be allocated to the inputs.

Definition 3. For a strictly monotone continuous generating function\(^3\) \( g : [0, 1] \to [-\infty, \infty] \) and weighting vector \( \mathbf{w} \), the weighted quasi-arithmetic mean is given by,

\[
QAM_{\mathbf{w}}(\mathbf{x}) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_i)\right). \tag{1}
\]

\(^3\)See [9] for more information regarding the choice of generators and their construction. Where \( g(0) \) or \( g(1) \) approach \( \pm\infty \), special care needs to be taken in calculation with the convention \( 0 \cdot \infty = 0 \) usually adopted. Methods also exist for using non-continuous and non-strict generators.
Special cases include weighted arithmetic means \( \sum_{i=1}^{n} w_i x_i \), where \( g(t) = t \), weighted power means \( (\sum_{i=1}^{n} w_i x_i^p)^{\frac{1}{p}} \), where \( g(t) = t^p \) and weighted geometric means \( G(x) = \prod_{i=1}^{n} x_i^{w_i} \) if \( g(t) = -\ln t \). The weights \( w_i \) are usually non-negative and sum to one.

Another well-known aggregation function is the ordered weighted averaging operator (OWA). Whereas the quasi-arithmetic means assign a weight to a particular input, the OWA first rearranges the inputs into descending order and then allocates importance accordingly.

**Definition 4.** For a weighting vector \( \mathbf{w} \), the ordered weighted averaging (OWA) operator is given by,

\[
OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)},
\]

(2)

where the parentheses \( . \) indicate a reordering of the inputs such that \( x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)} \).

Special cases include the maximum when \( \mathbf{w} = (1, 0, \ldots, 0) \), the minimum when \( \mathbf{w} = (0, \ldots, 0, 1) \) and the median if \( w_i = 1 \) for \( i = \frac{n+1}{2} \) and 0 otherwise where \( n \) is odd, and \( w_i = 0.5 \) for \( i = \frac{n}{2}, \frac{n}{2} + 1 \) and 0 otherwise where \( n \) is even.

The Choquet integral, defined with respect to a fuzzy measure, generalizes both the OWA and the weighted arithmetic mean. It is able to account for both individual contributions and the contributions of each subset or coalition.

**Definition 5.** Let \( \mathcal{N} = \{1, 2, \ldots, n\} \). A discrete fuzzy measure is a set function \( \mu : 2^{\mathcal{N}} \to [0, 1] \) which is monotonic (i.e. \( \mu(A) \leq \mu(B) \) whenever \( A \subset B \)) and satisfies \( \mu(\emptyset) = 0 \) and \( \mu(\mathcal{N}) = 1 \).

**Definition 6.** The discrete Choquet integral with respect to a fuzzy measure \( \mu \) is given by

\[
C_{\mu}(\mathbf{x}) = \sum_{i=1}^{n} x_{(i)}[\mu(\{j \mid x_j \geq x_{(i)}\}) - \mu(\{j \mid x_j \geq x_{(i+1)}\})],
\]

(3)
where \((x_1, x_2, \ldots, x_n)\) is a non-decreasing\(^4\) permutation of the input \(x\), and \(x_{n+1} = \infty\) by convention.

A number of indices have been introduced to measure concepts such as the overall importance of each input [13, 15]. The most commonly used of these is the Shapley index. It provides a set of \(n\) values which can be interpreted similarly to \(w\) for quasi-arithmetic means.

**Definition 7.** Let \(\mu\) be a fuzzy measure. The Shapley index for every \(i \in \mathbb{N}\) is
\[
\phi(i) = \sum_{A \subseteq \mathbb{N} \setminus \{i\}} \frac{(n - |A| - 1)!|A|!}{n!} \left[ \mu(A \cup \{i\}) - \mu(A) \right].
\]
The Shapley value is the vector \(\phi(\mu) = (\phi(1), \ldots, \phi(n))\). It satisfies \(\sum_{i=1}^{n} \phi(i) = 1\).

The last function we present is the Bonferroni mean [11] and its generalization [5]. The original definition is as follows.

**Definition 8.** Let \(p, q \geq 0\) and \(x_i \geq 0, i = 1, \ldots, n\). The Bonferroni mean is the function
\[
Bonf(x) = \left( \frac{1}{n(n-1)} \sum_{i,j=1, i \neq j}^{n} x_i^p x_j^q \right)^{\frac{1}{p+q}}. \tag{4}
\]

We see that the Bonferroni mean essentially takes the average of product pairs \(x_i x_j\), however it can also be rearranged such that we take the product of each input with the average of those remaining. In the case of \(p = q\) and \(n = 2\), the Bonferroni mean is equivalent to the geometric mean. As the ratio \(\frac{p}{q}\) approaches infinity (or 0), it approaches the maximum operator. It will always be required that at least two inputs are above zero (i.e. one of the products \(x_i x_j\)) in order to obtain a non-zero output. This property makes the Bonferroni mean and its generalizations an appropriate choice if it is desired to model mandatory requirements.

Its generalization is given in [5] and can be expressed as follows.

---

\(^4\)Note that this is opposite to the OWA, which is usually expressed in terms of a non-increasing permutation of the inputs.
Definition 9. Let $\mathbb{M}$ denote a 3-tuple of aggregation functions $(M_1, M_2, C)$, with $M_1 : [0, 1]^n \to [0, 1]$, $M_2 : [0, 1]^{n-1} \to [0, 1]$ and $C : [0, 1]^2 \to [0, 1]$, with the diagonal of $C$ denoted by $C_*(t) = C(t, t)$ and inverse diagonal $C_*^{-1}$. The generalized Bonferroni mean is given by,

$$B_{\mathbb{M}}(x) = C_*^{-1}\left(M_1\left(C\left(x_1, M_2(x_{j\neq 1})\right), \ldots, C\left(x_n, M_2(x_{j\neq n})\right)\right)\right),$$

with $x_{j\neq i}$ denoting the vector in $[0, 1]^{n-1}$ that includes the arguments from $x \in [0, 1]^n$ in each dimension except the $i$-th, $x_{j\neq i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

2.2. Learning aggregation weights using linear programming

From a given dataset of observed inputs and outputs, we can apply fitting procedures both for regression/classification and for data analysis, i.e. in our case, we can:

- (as a regression tool) learn aggregation weights or parameters from the data obtained in the online exercise (or from a selection of experts), and then use these parameters to predict or estimate the species diversity of new (unseen) communities;

- (as an analysis tool) use the dataset to make inferences about the importance of each of the variables in human judgements of species diversity, by learning and interpreting the weights of some common families of aggregation functions.

For the experiments in this work, we adopt the approach of optimizing the choice of weighting vector $w$ with respect to the least absolute deviation (LAD) of residuals [10, 1]. In the standard case for weight learning, we have a function $f_w$ and a set of observed values $y_k$ which we want the function to predict once we know its parameters. This leads to the following:

$$\text{Minimize } \sum_{k=1}^{K} |f_w(x_k) - y_k|.$$

Compared to a least squares approach, the least absolute deviation method is less sensitive to outliers and also can be formulated as a linear program.

---

5In order for the inverse diagonal to be well defined, $C_*$ should be injective. This can be ensured by choosing $C$ to be a strictly monotone and continuous T-norm or T-conorm.
which allows for a quick and efficient solution. Firstly, we denote by $r_k^+$ and $r_k^-$ the positive and negative parts of the difference $f_w(x_k) - y_k$. For each observed input/output pair $(x_{k1}, x_{k2}, \ldots, x_{kn}, y_k)$, one of the $r_k^+, r_k^-$ will be zero.

The weight learning can then be performed with the objective of minimizing the residuals with the following linear program.

\[
\begin{align*}
\text{Minimize} & \quad \sum_{k=1}^{K} (r_k^+ + r_k^-), \\
\text{s.t.} & \quad f_w(x_k) - r_k^+ + r_k^- = y_k, \quad k = 1 \ldots K \quad (7) \\
& \quad r_k^+, r_k^- \geq 0.
\end{align*}
\]

Weighted quasi-arithmetic means can still be fit linearly with appropriate data transformations, i.e. for a generating function $g$, we can use the constraints:

\[
\left( \sum_{i=1}^{n} w_i g(x_{ki}) \right) - r_k^+ + r_k^- = g(y_k), \quad k = 1 \ldots K
\]

\[
w_i \geq 0, \forall \ i, \quad \sum_{i=1}^{n} w_i = 1.
\]

Note that the residuals in this case are the differences between the generator transformed data (not the actual data itself).

For ordered functions such as the OWA, the data can be rearranged before the fitting procedure so that the weights are learned from the reordered data. In both cases, although the functions themselves are not linear, the weights are only fit to linear data. The fuzzy measure $\mu$ of the Choquet integral can similarly be fit in this way, however we note that for $n$-dimensional data, $\mu$ is defined by $2^n$ values (see [1] for detail on this process).

For the generalized Bonferroni mean, since $M_1$ is an averaging function of $n$ arguments while $M_2$ is a function of $(n - 1)$ arguments, they will have weighting vectors of different dimension. For the weighting vector associated with $M_2$, we employ a convention described in [5], whereby we use a weighting vector $u \in [0, 1]^n$ and for each $x_{j\neq i}$ associate an $(n - 1)$-dimensional vector $u^j$ with weights $u_j/(1 - u_i)$ for all $j$. We then can fit generalized Bonferroni
means by fixing the choice of means (and corresponding generators) and some of the weighting parameters (as in [3]). We consider two approaches for the experiments conducted in this work.

Firstly, we fix $M_1, M_2$ as weighted arithmetic means with respect to weighting vectors $w$ and $u$ respectively, and $C$ the product operation with powers $p = q = 1$. This leads to the following expression and simplification,

$$B_{M}(x) = \left( \sum_{i=1}^{n} w_i x_i \left( \sum_{j \neq i} \frac{u_j}{1 - u_i} x_j \right) \right) = \left( \sum_{i=1,j=1,i \neq j}^{n} \frac{w_i u_j}{1 - u_i} x_i x_j \right).$$

Although we still cannot separate the weights linearly, we can consider each $x_i x_j$ term and the coefficients $v_{ij} = w_i u_j / (1 - u_i)$. We hence transform the instances of the dataset $(x_{k1}, x_{k2}, \ldots, x_{kn}, y_k)$ such that we fit to a set consisting of all product pairs $(x_{k1} x_{k2}, x_{k1} x_{k3}, \ldots, x_{k1} (n-1) x_{kn}, y_k)$ and introduce the following linear constraints, minimizing with respect to the weights $v_{ij}$.

$$\left( \sum_{i=1,j=1,i \neq j}^{n} v_{ij} x_{ki} x_{kj} \right) - r_k^r + r_k^- = y_k, k = 1 \ldots K$$

$$v_{ij} \geq 0, \forall i, j,$$

$$\sum_{i=1}^{n} v_{ij} = 1.$$

The resulting $v_{ij}$ will not be separable into the $w_i, u_i, u_j$ etc, however these weights themselves represent the contribution of the pair $x_i x_j$ to the overall output. We can gain an idea of each input’s average contribution by summing the rows and columns of the $v_{ij}$ matrix respectively.

An alternative to fitting to the pairs $x_i x_j$ is to fix the weighting vector $u$. This way, we can use alternative means for $M_1, M_2$ (in particular, any QAM) whereas before we were limited to weighted arithmetic means. We hence perform fitting by transforming each of the input terms $x_{ki}$ by combining with the mean of the $x_{k,j \neq i}$ and using the generator functions. Denoting the generator of $M_1$ by $m_1$, each of the terms will be given by

$$m_1 \left( x_{ki} \left( M_2(x_{k,j \neq i}) \right) \right),$$

where the weighting vectors for $M_2$ are each of the $u^i$ determined from the supplied vector $u$. In this case, we introduce the following constraints and
fit with respect to the vector \( \mathbf{w} \).

\[
\left( \sum_{i=1,j=1,i\neq j}^{n} w_i m_1(x_{ki}(M_2(x_{k,j\neq i}))) \right) - r_k^+ + r_k^- = m_1(y_k),
\]

\[k = 1 \ldots K,
\]

\[w_i \geq 0, \quad \forall i,
\]

\[\sum_{i=1}^{n} w_i = 1.
\]

The weights fitted through this process can still be interpreted as the average importance of each \( x_i \) as they do for weighted quasi-arithmetic means, however it should be noted that since we take the conjunction with the remaining inputs, each weight will be applied to both \( x_i \) and those remaining (i.e. \( x_{j\neq i} \)), with \( w_i = 1 \) signifying that a high value for \( x_i \) is mandatory, but not sufficient, for a high output. In the results sections we will refer to Bonferroni means learned through this approach as \( \text{Bonf}_w \), whereas \( \text{Bonf}_{w_{ij}} \) will be used for the former method.

In [6] we presented a method for learning aggregation weights from comparisons of the form \( y_i > y_j \) rather than the \( y \) values themselves. In this case, we consider a target dataset \( \mathcal{D} \) containing \( K \) instances \( \mathbf{x}_k = (x_{k1}, \ldots, x_{kn}) \) and a set of pairwise comparisons \( \mathcal{P} \), where \((i,j) \in \mathcal{P}\) denotes the judgement that \( f_w(\mathbf{x}_i) \) should be greater than \( f_w(\mathbf{x}_j) \). We hence are looking for a function that satisfies \( f_w(\mathbf{x}_i) > f_w(\mathbf{x}_j) \), \( \forall (i,j) \in \mathcal{P} \) as much as possible.

This problem remains linear and can be formulated in the following way

Maximize

\[
\sum_{(i,j) \in \mathcal{P}} r_{ij}^+ - \lambda r_{ij}^-,
\]

s.t. \( f_w(\mathbf{x}_i) - f_w(\mathbf{x}_j) - r_{ij}^+ + r_{ij}^- = 0, \forall (i,j) \in \mathcal{P}, \)

\[
\sum_{j=1}^{n} w_j = 1,
\]

\[
r_{ij}^+, r_{ij}^- \geq 0, \forall (i,j) \in \mathcal{P},
\]

\[
w_j \geq 0, j = 1, \ldots, n.
\]

The \( \lambda \) value is a penalty parameter and can be used to control how much we allow the supplied pairwise judgements to be violated, i.e. each \( r_{ij}^- \neq 0 \) means that \( f_w(\mathbf{x}_i) - f_w(\mathbf{x}_j) < 0 \), even though we supplied the preference
that \( f_w(x_i) \) should be larger. If we wish only to minimize the number of violations, then we can remove the \( r_{ij}^+ \) terms from the objective function, however including it the way it is here means that if a comparison is given, then we want the separation between the scores for \( x_i \) and \( x_j \) to be as large as possible. In [6], we ran experiments to identify weights based on comparisons drawn from a synthetically generated dataset. We found that \( \lambda = 5 \) is an appropriate value, which doesn’t overfit the data but also allows a reasonable level of accuracy. For the experiments in this work, we also used the constraint,

\[
f_w(x_i) - f_w(x_j) - r_{ij}^+ + r_{ij}^- = 0.00001
\]

in the fitting algorithms rather than using 0 on the right hand side. This is so that we have at least this level of strict inequality whenever a comparison pair is satisfied.

We can still use generator transformations to fit the weights of quasi-arithmetic means with this method, however we need to ensure that the generating function is increasing.

3. Intuitive human judgements of species diversity

In this section we will firstly provide some background on ecological indices used for measuring species diversity, followed by a description of the real dataset we have obtained and which will be used for the regression and analysis experiments performed in the following sections.

3.1. Species Diversity and Ecological Indices

In environmental ecology the concept of species diversity incorporates both the species richness of a community and the relative abundance of each species. It is usually held (and argued in [22]) that this relationship should be multiplicative, i.e., of the form \( D = E \times R \). Richness (\( R \)) refers to the raw number of species (at the level of measurement) and so there is little debate around how it should be measured, however more and more indices have been proposed for evenness (\( E \)). The behavior of many of these evenness indices has been investigated in [20] and [22] (among others) and we have looked at their properties (in the context of consensus measures), particularly in terms of monotonicity and limiting behavior, in [7, 8].

In this paper, we will focus on two of the most commonly used and simplest expressions for diversity. For an ecological community consisting of
n species (i.e. \( R = n \)), the proportional abundance of the \( i \)-th species is given by \( p_i \). The species diversity of that community can then be calculated using

\[
D_1(p) = \frac{1}{\sum_{i=1}^{n} p_i^2},
\]

or

\[
D_2(p) = \frac{n \left( 1 - \sum_{i=1}^{n} p_i^2 \right)}{1 - 1/n}.
\]

The sum of squared species proportions is known as Simpson’s dominance index [19] and has its roots in the study of equitability in economics. We also note that sometimes the terms \( p_i \ln p_i \) are used instead, providing an analogous equation to Shannon’s information entropy [18]. The key behavior captured by these calculations is that the overall species dominance is reduced as proportional abundance is transferred from a more abundant to a less abundant species. The value of \( D_1 \) can be interpreted as the effective number of species present in a population. When the proportional abundances are equal, \( D_1 \) will be the same as the value for richness. When only one species is present, it will give its minimum output which is 1. The output for \( D_2 \) on the other hand ranges from 0 to \( n \), reaching a maximum of \( n \) when all species are equal in abundance and zero when a single species dominates.

We will use these biodiversity indices as a benchmark for our aggregation frameworks for evaluating diversity, the weights of which will be learnt from the following dataset.

3.2. Our dataset

The dataset we will use for our investigation consists of 48 sets of responses to an online exercise conducted in 2014. We created a set of 252 hypothetical communities with richness values of 2, 3, 4, 5, 6, 8 and 10. For each value of richness, various species abundance distributions were generated randomly with total abundances of 100, 200 and 300 individuals.

The distributions were presented as charts (e.g. see Figs. 1 - 2) and included approximately uniform distributions (Fig 1(a)), those with a concave or convex increase in the number of species (by ‘concave’ we mean the difference between the least abundant species and second least abundant species is higher than the difference between the most abundant and second most abundant species, then the reverse for convex, see Figs. 1(b)-(c)), and
Figure 1: Examples of randomly generated species abundances for total abundance of 300 and richness of 5 according to desired type. The red lines are intended only to emphasize the distribution shape and were not shown in the survey.
threshold situations where the number of species jumped suddenly at the $i$-th species (Figs. 1(e)-(f)). In the Fig. 1 examples, all communities have the same richness and same overall abundance, however we might think that the community in (f) is less diverse because essentially there are only two species that have significant abundances, whereas in (a) each of the species is equally abundant.

We constructed two surveys consisting of 16 questions where participants were asked to make intuitive comparisons of three community charts at a time in terms of their diversity and conservation importance. An example of one of these questions and the online layout is shown in Fig. 2.

![Figure 2: One of the questions from the online exercise. Participants select the rankings from drop-down boxes to indicate their ordering.](image)

Each question was set up to include a base image (not necessarily the left image) along with two others that differed only in terms of distribution, richness or abundance. For example, in the pictured question all three communities have an abundance of 200, images 1 and 2 have the same richness and images 1 and 3 have the same distribution (concave with increasing species). For the purposes of our current experiments, we focus only on the diversity rankings given by the participants to the first 10 questions, yielding up to
30 comparison pairs\textsuperscript{6,7}. We were able to obtain 32 and 16 response sets (48 overall) to the two respective surveys, giving us a total of 1083 comparison pairs.

We note that although 48 participants is a relatively small sample size for drawing conclusions about the wider population, the models we use have only a small number of free parameters and the data points we actually fit to are the comparison pairs (rather than an individual). In practice, the number of experts we have access to for building such models may also be limited, however we can still use methods such as cross evaluation and leave-one-out testing in order to assess the robustness and reasonableness of the derived parameters. Our results support the notion that learning weighted aggregation functions from a set of expert opinions may produce biodiversity evaluations that are closer to human intuition than any one index alone.

4. Fitting aggregation functions to quantify diversity

Here we evaluate the ability of aggregation functions to model the species diversity evaluations given by participants in the online exercise.

4.1. Alignment between ecological indices and participant responses

Firstly, we will assess the correspondence between the participants’ evaluations and the basic indices of abundance, richness, evenness and diversity that are used in ecology. Table 1 shows the accuracy for each of the individual indices in predicting the participants’ intuitive comparisons of the hypothetical communities, i.e. if the participant judged chart A to show a more diverse community than chart B, then it is required that the index relating to A should also be greater than that for chart B. Abundance refers to the total abundance (rather than individual abundances), while $E_1$ and $E_2$ are the evenness calculations corresponding to $D_1$ and $D_2$ (and can be calculated by dividing the diversity indices by the richness $n$).

From the 16 respondents to Survey 1, we have a total of 391 comparison pairs and we obtained a total of 792 from the 32 respondents to Survey 2.

\textsuperscript{6}The final 6 questions included similar community triples to the first six questions, but also indicated the presence of either a rare or keystone species. We do not consider these variables and questions in the current work.

\textsuperscript{7}Less than 30 preference pairs were obtained if participants skipped questions or allocated two communities an equal rank.
No comparisons were generated if the respondent ranked two communities to be equal in terms of diversity or if the respondent did not answer that particular question.

Table 1: Percentage of community pairs where the index accurately reflects the participants’ judgements

<table>
<thead>
<tr>
<th></th>
<th>Abundance</th>
<th>Richness</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$D_1$</th>
<th>$D_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey 1 (16)</td>
<td>19.9</td>
<td>50.6</td>
<td>60.9</td>
<td>52.2</td>
<td>77.2</td>
<td>79.5</td>
</tr>
<tr>
<td>Survey 2 (32)</td>
<td>25.0</td>
<td>44.4</td>
<td>42.3</td>
<td>35.1</td>
<td>65.7</td>
<td>68.7</td>
</tr>
</tbody>
</table>

The poor performance of abundance and richness is largely due to the limited granularity in terms of the possible responses, i.e. for abundance there are only 3 options and 7 options for richness. Although one would expect richness to be an important consideration for the human judgement process, obviously it is not able to compare two communities with the same number of species present. The diversity indices on the other hand, which incorporate both richness and evenness correspond quite well with the majority of participant evaluations.

4.2. Using aggregation functions to predict participant responses

We are interested in whether some other aggregations of abundance, richness and evenness might better capture the behavior in the respondents’ evaluations, or additionally, whether aggregations involving diversity along with the other indices might produce better results than any of the indices alone. We hence conducted two sets of experiments whereby the weights of different aggregation functions were learnt using the linear fitting method in Eq. (8), and then the resulting functions were used to predict the responses of the each participant.

4.2.1. Leave-one-out approach for predicting responses to the same survey

The results to the first set of experiments is shown in Table 2, where a leave-one-out train/test method was followed (bold denotes best (or equal best) performing function with respect to the dataset used). Each community is represented by the set of indicated indices, i.e. we tried fitting the aggregation functions to 3-variable data, consisting of abundance, richness and either of the evenness indices, as well as to 6-variate data where each community is described by its scores for all six ecological indices. As an example, for fitting to the indices $A, R, E_1$, the three communities from Fig.
2 would be associated with the vectors \( \mathbf{x}_1 = (0.666667, 0.4, 0.950133), \mathbf{x}_2 = (0.666667, 0.4, 0.950133), \mathbf{x}_3 = (0.666667, 0.6, 0.984433) \), each of the three indices being scaled to take values in \([0, 1]\).

Table 2: Percentage of comparisons accurately predicted using various aggregation functions. The weights are learned from the responses provided by all the other respondents to the same survey.

<table>
<thead>
<tr>
<th>Indices used</th>
<th>Survey 1</th>
<th>Survey 2</th>
<th>Survey 1</th>
<th>Survey 2</th>
<th>Survey 1</th>
<th>Survey 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( A, R, E_1 )</td>
<td>( A, R, E_2 )</td>
<td>( A, R, E_1, E_2, D_1, D_2 )</td>
<td>( A, R, E_1, E_2, D_1, D_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( A, R, E_1 )</td>
<td>( A, R, E_2 )</td>
<td>( A, R, E_1, E_2, D_1, D_2 )</td>
<td>( A, R, E_1, E_2, D_1, D_2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AM</td>
<td>80.1</td>
<td>65.5</td>
<td>84.1</td>
<td>80.7</td>
<td>77.2</td>
<td>68.7</td>
</tr>
<tr>
<td>GM</td>
<td>19.9</td>
<td>42.3</td>
<td>56.0</td>
<td>35.1</td>
<td>56.0</td>
<td>42.3</td>
</tr>
<tr>
<td>QM</td>
<td>82.4</td>
<td>65.5</td>
<td>84.1</td>
<td>80.7</td>
<td>77.2</td>
<td>68.7</td>
</tr>
<tr>
<td>OWA</td>
<td>82.1</td>
<td>73.5</td>
<td>80.1</td>
<td>80.7</td>
<td>77.2</td>
<td>68.7</td>
</tr>
<tr>
<td>( Bonf_w )</td>
<td>84.4</td>
<td>80.7</td>
<td>81.3</td>
<td>79.2</td>
<td>86.4</td>
<td>81.6</td>
</tr>
<tr>
<td>( Bonf_{v_j} )</td>
<td>84.1</td>
<td>79.3</td>
<td>84.1</td>
<td>79.2</td>
<td>79.5</td>
<td>67.8</td>
</tr>
<tr>
<td>( Ch )</td>
<td>83.6</td>
<td>66.8</td>
<td>80.8</td>
<td>72.0</td>
<td>77.2</td>
<td>68.7</td>
</tr>
</tbody>
</table>

In this case, to predict the judgements of each participant, we use the combined set of comparisons from all other respondents to the same survey as the training set. We see that even taking an aggregation of abundance, richness and evenness (i.e. not including the existing diversity indices) is able to predict the participants’ judgements more accurately than the diversity indices (with the exception of the geometric mean which perhaps does not perform well since it is affected more by increases to smaller inputs than larger inputs). However of course in this case, the test set uses the same set of comparisons over which the training set is defined and the flexibility of the aggregation functions would allow them to achieve more accuracy, especially if respondents are more or less similar in their responses.

4.2.2. Fitting aggregation functions to one survey in order to predict responses to the other

The results in Table 3 are obtained by using the responses from Survey 1 in order to predict the responses to Survey 2 and vice-versa.

The surveys were designed such that as few communities as possible were replicated in the two sets of 30 that were shown to participants, so it should not be the case that these results are biased by the comparisons used. Once again, the results here show a significant improvement to those obtained from
Table 3: Accuracy for various aggregation functions where weights are learned from the responses from Survey 1 were used to predict those from Survey 2 and vice versa.

<table>
<thead>
<tr>
<th>Indices used</th>
<th>A, R, E₁</th>
<th>A, R, E₂</th>
<th>A, R, E₁, E₂, D₁, D₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>AM</td>
<td>60.6</td>
<td>76.3</td>
<td>85.2</td>
</tr>
<tr>
<td>GM</td>
<td>60.9</td>
<td>25.0</td>
<td>52.2</td>
</tr>
<tr>
<td>QM</td>
<td>60.6</td>
<td>76.3</td>
<td>85.2</td>
</tr>
<tr>
<td>OWA</td>
<td>70.1</td>
<td>77.4</td>
<td>70.1</td>
</tr>
<tr>
<td>Bonf_w</td>
<td>85.2</td>
<td>80.7</td>
<td>81.3</td>
</tr>
<tr>
<td>Bonf_vij</td>
<td>84.1</td>
<td>76.3</td>
<td>84.1</td>
</tr>
<tr>
<td>Ch</td>
<td>68.5</td>
<td>76.3</td>
<td>68.5</td>
</tr>
</tbody>
</table>

using diversity by itself. A potential explanation for this is that the diversity indices essentially allow a tradeoff between evenness and richness. The aggregation functions used here on the variables A, R, E₁ and A, R, E₂ are able to modify the extent of this tradeoff according to the intuitive human perceptions provided in the dataset, as well as include the overall abundance in the evaluation. Where the diversity calculations are included in the set of indices, clearly here the aggregation functions have the flexibility to compensate between D₁ and D₂, as well as the other variables in order to provide a more robust prediction.

We make special note of the Bonferroni mean, which has the best overall performance. The Bonferroni mean’s ability to model mandatory requirements seems especially suited here, since effectively it can ensure that both richness and evenness are high in order to give a high score for diversity, but without overfitting in the way that the Choquet integral does. It is also worth noting that the geometric mean often performed poorly, even though a weighting vector that allocates equal importance to evenness and richness (with zero to abundance) should induce the same comparisons as those that would be obtained with the diversity indices, i.e. in the fitting procedure we have \( R = n/10 \) and hence a geometric mean \( GM(x) = \sqrt{x_2x_3} = \sqrt{ER/10} \) is simply the square root of the corresponding diversity calculation divided by the constant \( \sqrt{10} \). Clearly the fitting procedure with the obtained pairs does not result in the required weighting vector \((0, 0.5, 0.5)\), perhaps being biased too much towards the supplied comparisons to perform well on those for prediction.
4.2.3. Fitting aggregation functions to raw abundances data

In Table 4 we show the results of fitting 10-variate functions to the individual species abundance data. In this case, rather than use the summary indices, we use the vector of species abundances, with 0 denoting a species being absent. For the previously used examples from Fig. 2, the input vectors would be \( x_1 = (0, \ldots, 0, 18, 52, 62, 68), x_2 = (0, \ldots, 0, 24, 40, 62, 74), x_3 = (0, \ldots, 0, 14, 31, 37, 39, 42). \) Since the species are in increasing order, we do not consider the OWA function or Choquet integral (they will be the same as the arithmetic mean) and nor do we consider the geometric mean since the ‘absent’ species values of 0 will always result in a 0 output. The case for aggregation functions used in this fashion is that an increase to any species should not result in a decrease to the overall diversity. The output generated for these functions could be interpreted as a kind of average abundance. It is not necessary for the inputs to be scaled in this case either.

Table 4: Accuracy for 10-variate aggregation functions various aggregation functions based on raw abundances.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Train data Same survey (leave-1-out)</th>
<th>Other survey responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>AM</td>
<td>83.6</td>
<td>81.9</td>
</tr>
<tr>
<td>QM</td>
<td>82.9</td>
<td>81.6</td>
</tr>
<tr>
<td>Bonf_w</td>
<td>66.2</td>
<td>81.7</td>
</tr>
<tr>
<td>Bonf_v</td>
<td>81.3</td>
<td>81.6</td>
</tr>
</tbody>
</table>

The dominance index used in the evenness evaluations is an aggregation function of this form, however with the weights equal to the inputs. We note that such weighting vectors would hence increase as the inputs increase. We would expect a vector generated using our learning methods to be the reverse of this and hence have higher weights for the lesser abundant species in order to favor communities with higher richness. For example, an OWA operator with respect to a weighting vector which doubles for each species (i.e. the second most abundant species has a weight double the first, the third most abundant species has double the second and so on) would always give a higher diversity value as the richness increases. Although the results are not shown for it in the table, we note that such a (pre-assigned) weighting vector would accurately predict 82.4 and 78.2 percent of the respondents’ comparisons for Survey 1 and 2 respectively. Presumably the fitting procedure is not able to
achieve this level of accuracy when one set of responses is used to predict the other due to overfitting.

We have shown that aggregation functions, in particular the Bonferroni mean, can be useful in modeling human judgements and predicting the relative species diversity of ecological communities. This suggests that in making intuitive comparisons, respondents perhaps think slightly differently about what diversity means or take into account subtle features of a community’s (numerical) description than is expressed by functions such as $D_1$ and $D_2$. In the following section, we use our fitting approach in order to uncover the implied importance of each of these features and perhaps shed some light on the results we have obtained thus far.

5. Using aggregation functions for analysis of human preferences

For the following analyses, we use the entire set of comparisons collected from the first 10 questions of the online exercise in order to interpret the overall behavior of the respondents. For each of the aggregation functions used in the previous section for prediction, we can now learn the associated weighting vectors that minimize the extent and number of violations with respect to the entire dataset of participant responses. We can interpret the resulting weights as being indicative of the importance or influence of each of the variables.

5.1. Aggregations of abundance, richness and evenness

Table 5 shows the results when the aggregation functions are fit to three indices: abundance, richness and evenness.

We see that the functions that performed well almost allocated their entire weight to richness, with small amounts of weight from abundance and evenness helping to distinguish between cases where the richness is the same. For the Bonferroni mean fit with respect to a fixed secondary weighting vector, all the weight was allocated to richness. This means the function has the form,

$$Bonf_w(x_1, x_2, x_3) = \sqrt{x_2 \left(\frac{x_1 + x_3}{2}\right)}.$$

This models richness as a mandatory criterion, but high richness alone is not enough to indicate high diversity if both abundance and evenness are low.
Table 5: Fitted weights indicating relative importance of indices when the entire set of respondent comparisons is used.

<table>
<thead>
<tr>
<th></th>
<th>AM</th>
<th>GM</th>
<th>QM</th>
<th>Bonf w</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A, R, E₁</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abundance</td>
<td>0.0003</td>
<td>0.00311</td>
<td>0.00003</td>
<td>0</td>
<td>0.00180</td>
</tr>
<tr>
<td>Richness</td>
<td>0.99995</td>
<td>0</td>
<td>0.99992</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>$E₁$</td>
<td>0.00002</td>
<td>0.99689</td>
<td>0.00005</td>
<td>0</td>
<td>0.49820</td>
</tr>
<tr>
<td>Fitting Acc</td>
<td>63.7</td>
<td>36.9</td>
<td>68.6</td>
<td>82.2</td>
<td>67.8</td>
</tr>
<tr>
<td><strong>A, R, E₂</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abundance</td>
<td>0.00003</td>
<td>0.01445</td>
<td>0.00003</td>
<td>0</td>
<td>0.02744</td>
</tr>
<tr>
<td>Richness</td>
<td>0.99992</td>
<td>0</td>
<td>0.99992</td>
<td>1</td>
<td>0.74147</td>
</tr>
<tr>
<td>$E₂$</td>
<td>0.00005</td>
<td>0.98555</td>
<td>0.00005</td>
<td>0</td>
<td>0.23108</td>
</tr>
<tr>
<td>Fitting Acc</td>
<td>70.6</td>
<td>46.7</td>
<td>73.4</td>
<td>80.9</td>
<td>72.9</td>
</tr>
</tbody>
</table>

For fitting a Bonferroni mean with weights $v_{ij}$, only the weights $v_{12}$ and $v_{23}$ were allocated any weight in both cases. For the set of indices with $E₁$ these weights were 0.02504 and 0.97496 respectively, while for the set of three indices with $E₂$, these values were 0.10614 and 0.89386. The fitting accuracy was 80.9 in both cases. As with $Bonf_w$ the output can only be high if both richness and evenness are high.

The weights for the Choquet integral shown in the table are actually those of the Shapley value, giving an overall idea of the importance for each index. For the set of indices using $E₁$, the values of the corresponding fuzzy measure were all zero except for $\mu(\{1, 2\}) = 0.00360$ and $\mu(\{2, 3\}) = 0.99640$. This means the function behaves similarly to $\min(x_2, x_3)$, i.e. just taking the minimum of richness and evenness. For the set of indices using $E₂$, $\mu(\{1\}) = 0$ and $\mu(\{3\}) = \mu(\{1, 3\}) = 0.00003$. We then have $\mu(\{2\}) = 0.48297, \mu(\{1, 2\}) = 0.53785, \mu(\{2, 3\}) = 0.94512$. This shows that richness and evenness together are quite important with less weight being allocated to evenness if it is higher than richness in value.

The weighting vectors for the OWA are not shown in the table since they are not interpreted with respect to the actual variables, however we note that these weights were $w = (0.00011, 0.33325, 0.66663)$ and $w = (0.33571, 0.08983, 0.57445)$ respectively - in both cases allocating more importance to the lowest value and behaving similar to a minimum function.
The fitting accuracy for the OWA was 76.2 and 80.1 respectively.

5.2. Aggregations that include existing diversity indices

We also performed fitting on the set of six indices in order to capture which of the indices, including diversity, are able to best capture the judgements of the participants. The weights of the fitted functions are shown in Table 6.

Table 6: Fitted weights indicating relative importance of indices when the entire set of comparisons is used with all six ecology indices.

<table>
<thead>
<tr>
<th></th>
<th>AM</th>
<th>GM</th>
<th>QM</th>
<th>Bonf_w</th>
<th>Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abundance</td>
<td>0</td>
<td>0.01445</td>
<td>0</td>
<td>0</td>
<td>0.26085</td>
</tr>
<tr>
<td>Richness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E₁</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E₂</td>
<td>0</td>
<td>0.98555</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D₁</td>
<td>0.21746</td>
<td>0</td>
<td>0</td>
<td>0.28962</td>
<td>0.47831</td>
</tr>
<tr>
<td>D₂</td>
<td>0.78254</td>
<td>0</td>
<td>1</td>
<td>0.71038</td>
<td>0.26085</td>
</tr>
<tr>
<td>Fitting Acc</td>
<td>71.7</td>
<td>46.7</td>
<td>72.3</td>
<td>83.2</td>
<td>71.7</td>
</tr>
</tbody>
</table>

The well performing functions all used a combination of D₁ and D₂ with the majority of the weight to D₂. This supports the idea that D₂ is quite capable of capturing the behavior of the respondents, however it also shows that a combination of both indices, aggregated together, is perhaps better than D₂ alone.

The OWA in this case had the same fitting accuracy as the quadratic mean, and allocated almost all weight to the highest input with \( w = (0.99090, 0.00910, 0, 0, 0) \). This is somewhat surprising given the opposite tendency for the sets of 3 indices.

For the Bonferroni mean with respect to weighted \( v_{ij} \) pairs, all the weight was allocated to \( v_{36} \) making the fitted function equivalent to the geometric mean of D₁ and D₂. The question of why the geometric mean was not able to capture this behavior, allocating the same weights and resulting in the same function may be attributable to transformations used and the fitting criteria. For the geometric mean, the transformation \( g = -\ln t \) is used and the functions are fit to this, whereas the terms in the Bonferroni mean fitting method are fit to each of the pairs \( x_i x_j \). For values less than 1, taking the natural logarithm can result in exponentially large differences between low values that are quite close before the transformation.
The fitted weights of the Choquet integral are such that the function behaves similarly to taking the maximum out of $A, D_1, D_2$ with slightly more weight to $D_1$.

Overall, we can see that the diversity indices commonly used in ecology correspond quite well with human judgements, however a combination of the two (or indeed, incorporating aspects of all the indices as $Bonf_w$ does) was able to more closely model the data. Both $D_1$ and $D_2$ essentially model the number of species, traded off against the distribution of their abundances. The subtle difference hence lies in the degree to which these features compensate for one another, e.g. for what kind of distribution a community with 5 species seems more diverse than a species with 6 species, and it is here that the weights learning process is able to capture the human preferences in a way that might not be possible a priori.

6. Discussion and Future Work

From these results, we see the potential of results in the field of aggregation functions to contribute to other sciences. In particular, the work in aggregation of fuzzy sets, for linguistic terms and in modeling human reasoning is well suited to sciences where it is necessary to quantify concepts that reflect human intuitions. For real uptake of methods such as is presented here, however, both the performance and ease of interpretation need to be established. We hope that this work makes a step in this direction. We mention the following lines of research that may extend the results found here.

6.1. Further use of the comparisons dataset

The dataset we obtained from participants also includes responses to questions involving the presence of a ‘rare’ species or a ‘keystone’ species (a species which influences its environment much more than would be expected from its numbers). It could be that such factors, along with the identification of the actual species, has an influence on how people actually perceive the diversity of an ecological community. In addition to using results from aggregation functions as was undertaken here, the difference between species can also be factored into a number of existing ecological indices (an observation we made in [8]).
The area of consensus is also an area that may benefit from use of a human participants survey, since consensus (like evenness) can also be considered as a somewhat fuzzy concept.

6.2. Fitting comparison data of different dimension

In this article we used an adaption of the least absolute deviation fitting method for dealing with comparison data rather than numerical assessments. In fitting to the raw abundances, we used vectors of consistent length (10) and used zeros to populate the vectors when the number of species was smaller. Alternatively, rather than attempting to find an overall vector \( \mathbf{w} \), we could also attempt to find multiple vectors of different dimension \( \mathbf{w}^{(2)}, \mathbf{w}^{(3)}, \ldots, \mathbf{w}^{(10)} \). We do not want these vectors to be completely independent, however. Ideally, we would construct a family of weighting vectors that are stable [17, 4] where the ratios between the weights \( w_1/w_2, w_2/w_3 \text{ etc.}, \) are the same for all weighting vectors. However this cannot be achieved linearly with respect to the weights. One possibility to make them approximately stable would be to define the weighting vectors such that

\[
\begin{align*}
    w_1^{(2)} &> w_2^{(2)} \iff w_1^{(3)} > w_2^{(3)}, \\
    w_2^{(3)} &> w_3^{(3)} \iff w_2^{(4)} > w_3^{(4)}, \ldots
\end{align*}
\]

and so on for all pairs of weighting vectors. This gives us either/or constraints, i.e.

Either

\[
\begin{align*}
    w_1^{(2)} &> w_2^{(2)}, \\
    w_1^{(3)} &> w_2^{(3)}, \\
    \ldots
\end{align*}
\]

Or

\[
\begin{align*}
    w_1^{(2)} &< w_2^{(2)}, \\
    w_1^{(3)} &< w_2^{(3)}, \\
    \ldots
\end{align*}
\]

and so on for all pairs \( w_i, w_j \).

Such constraints can be modeled and solved by introducing a binary variable \( t \), which, although requiring integer programming methods, allows the problem to remain linear. This is achieved by including sets of constraints

25
of the following form into our weight fitting algorithm,

\[
\begin{align*}
    w_1^{(2)} - w_2^{(2)} + tM & > 0 \\
    w_2^{(2)} - w_1^{(2)} + (1 - t)M & > 0 \\
    \cdots \\
    w_1^{(10)} - w_2^{(10)} + tM & > 0 \\
    w_2^{(10)} - w_1^{(10)} + (1 - t)M & > 0
\end{align*}
\]

where \( t \) is binary and \( M \) is an arbitrarily large constant.

The increased complexity of the problem and the resulting number of variables may mean such approaches for large \( n \) are not particularly useful, however definitely hold potential where vectors differing in dimension by one to a few variables need to be found.

7. Conclusions

We have applied aggregation function fitting techniques to the field of ecological diversity indices. We did this with two intentions:

- To assess the extent to which current environmental indices used to quantify diversity reflect intuitive evaluations;
- To show the potential of aggregation functions to provide a mathematical framework for quantifying biodiversity that models the way humans think of diversity in ecological environments.

Overall, we found that the generalized Bonferroni mean was particularly suited to this dataset, since it is able to model mandatory requirements. The concept of diversity in ecology is one that is mainly dependent on richness, i.e. the number of species present. However in some cases the distribution of species may be such that the perception of how many species there effectively are is reduced, e.g. if one species makes up 99% of the population.

The Bonferroni mean is able to require both richness and the evenness of the distribution to be high, as well as some other contributing variables, in order to produce a high diversity evaluation. In this case, this behavior was better able to model the dataset than simply taking a weighted average (where a high value in richness is able to compensate for a low evenness).

To undertake this work, we also created a new dataset that may be of use to the soft computing and machine learning community and further developed techniques for fitting aggregation functions to data.
Acknowledgement

This work was supported by the Deakin University Central Research Grants Scheme, project number RM25299.

References


