ELSEVIER

Indices for monitoring biodiversity change: Are some more effective than others?

Eric G. Lamb [a,*], Erin Bayne [a], Gillian Holloway [a], Jim Schieck [b], Stan Boutin [a], Jim Herbers [a], Diane L. Haughland [a]

[a] Alberta Biodiversity Monitoring Institute, Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9
[b] Alberta Biodiversity Monitoring Institute, Alberta Research Council, PO Bag 4000, Vegreville, Alberta, Canada T9C 1T4

ARTICLE INFO

ABSTRACT

Statistically rigorous methods for summarizing and reporting trends in the intactness of biodiversity are a key element of effective biodiversity monitoring programs. There are four major approaches for translating complex monitoring data into easily communicated summary statistics: (1) traditional diversity indices such as species richness and Simpson's diversity, (2) species intactness indices based on occurrence, (3) species intactness indices based on abundance, and (4) multivariate community indices. We use simulated data to evaluate the effectiveness of 13 indices from these four categories based on statistical robustness, sensitivity to errors and noise in the data, ecological relevance, and ease of communication. We show that indices that calculate species intactness using equations like Buckland's arithmetic mean index are the most effective for use in large-scale biodiversity intactness monitoring programs. Traditional diversity indices are unsuitable for monitoring of biodiversity intactness, and multivariate indices can be highly sensitive to errors and noise in the data. Finally, we provide guidelines for the application of these indices in biodiversity intactness monitoring.

1. Introduction

Monitoring is a critical element in efforts to conserve and manage biodiversity. A successful monitoring program must use statistically rigorous methods to evaluate changes in biodiversity intactness over time (Noss, 1990; Debinski and Humphrey, 1997; Yoccoz et al., 2001; Nichols and Williams, 2006; Lovett et al., 2007). Effectively summarizing and communicating results from biodiversity monitoring to a broad diversity of people including the scientific community, government, industrial stakeholders, and the general public is challenging however due to the complexity of the data and the diverse needs of users. As such, most users and many policy decisions rely heavily on indices that compile and summarize monitoring data across large numbers of species (Overton et al., 2002). Indices summarizing trends in the intactness of biodiversity, or the degree to which observed diversity deviates from natural or reference conditions (hereafter intactness), are important to effective communication. In this paper we evaluate four major approaches for translating complex monitoring data into summary statistics: (1) traditional diversity indices, (2) species intactness indices based on occurrence, (3) species intactness indices based on abundance, and (4) multivariate community intactness indices.

* Corresponding author. Current address: Department of Soil Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 5A8. Tel.: +1 306 966 6855; fax: +1 306 966 6881.
E-mail address: eric.lamb@usask.ca (E.G. Lamb).

Traditional biodiversity metrics rely on species counts and or composite indices such as Shannon's or Simpson's index (Simpson, 1949; Margalef, 1958). Widely used in ecology, such indices have the benefit of being relatively easy to understand, simple to calculate, and have a long history of application. Richness and diversity are important ecological state variables, but retain only a small portion of the available information that describes the concept of biodiversity (Magurran, 2004). For example, the establishment of a non-native invasive species will increase the value of standard diversity indices. While more species may be present, most ecologists would argue that such a change is not positive for the maintenance of native biodiversity intactness. Thus, providing a manager a reporting tool based on a traditional diversity index should be done cautiously as knowledge of why a diversity index goes up or down is critical for proper interpretation.

To avoid the risks in interpretation caused by using traditional diversity indices, recent approaches have tried to measure the magnitude of change of individual species relative to some reference state (e.g. Buckland et al., 2005; Loh et al., 2005; Nielsen et al., 2007). These indices measure overall biodiversity intactness by summarizing the intactness of individual species or the average intactness of a group of species. Such indices reduce the possibility of misinterpretation because the intactness of each species is measured in a common metric that can be applied to measure the overall intactness of all the species in a defined area. In such indices, individual identity is maintained and by comparing observed data for each species to a defined reference state "positive" vs. "negative" changes in biodiversity intactness can be identified. Benchmarks can be derived through a variety of methods including expert identification of a desired goal or target (e.g. Taft et al., 2006), the earliest point on a time series (e.g. Loh et al., 2005), protected areas (e.g. Sinclair et al., 2002; Scholes and Biggs, 2005), or empirically derived reference conditions (Nielsen et al., 2007).

Two categories of species intactness indices are commonly used (Nielsen et al., 2007). Species intactness indices based on occurrence utilize the proportion of sites where a species was detected while abundance indices utilize some measure of species abundance. Combinations of both occurrence and abundance indices have also been proposed to calculate comprehensive species intactness indices that separate the factors that influence trends in abundance from occurrence patterns (Nielsen et al., 2007).

Biodiversity intactness can also be monitored at the community level through multivariate approaches. Communities are characterized by predictable combinations of species identity and abundance (Chapin et al., 2000), and deviations of the observed patterns from a reference pattern can be used as a measure of biodiversity intactness. Examining intactness at the level of the community provides a natural way to group species that belong to different trophic levels or taxonomic groups into a single index. When species intactness indices are averaged across species belonging to more than one taxonomic group the more species-rich group will dominate the result. Thus changes in the biodiversity intactness of the less species-rich group may be obscured by the larger group. Multivariate indices of diversity intactness directly deal with differences in richness between taxonomic

groups or trophic levels because those differences are a part of the multivariate patterns of species identity and abundance that characterize that community.

In this paper we test the effectiveness of three traditional diversity indices, three species intactness indices based on occurrence, four species intactness indices based on abundance and three multivariate community indices. Three criteria were used to rank the effectiveness of these indices: (1) statistical robustness was examined by describing the statistical properties of each index, including the power to detect known trends, and sensitivity to detection error, (2) ecological relevance was assessed for each index by measuring how the index performed in a series of ecological scenarios where known numbers of species invaded or were extirpated from communities, and (3) the ease of communication of each index based on the mathematical complexity of the index and the simplicity of reporting results to scientists, policy makers, and the general public.

1.1. Traditional indices

(1) *Species richness*: this is the simplest measure of diversity and was defined as the number of the species found in a defined area. Composite diversity indices that incorporate both species richness and evenness are also commonly used. Two of the most widely used indices are Shannon's index, H', and Simpson's index.

(2) *Shannon index* (Margalef, 1958):

$$H' = -\sum_i p_i \log(p_i)$$

where s is the number of species and p_i is the proportion of the sample belonging to the ith species.

(3) *Simpson index* (Simpson, 1949):

$$D = \sum_i (p_i)^2$$

where p_i is the proportion of the sample belonging to the ith species. Since D is the probability that two random individuals belong to the same species, Simpson's index of diversity is generally calculated as the complement of D:

$$1 - D = 1 - \sum_i (p_i)^2$$

We have used Simpson's index for infinite populations, since the corrected formula for a finite population is only appropriate for count data. Many species measured in a biodiversity monitoring program cannot be counted and must be measured using percent cover or biomass estimates which cannot be used in the finite population formula (Krebs, 1989). Traditional measures of diversity do not directly provide information on how observed diversity differs from benchmark or reference conditions. In this paper we rescaled the richness and diversity measures to include benchmark conditions using the following formula:

$$\text{diversity intactness index} = \left(\left(\frac{R - O}{R} \right) + 1 \right) \times 100$$

where R is the reference diversity calculated on the benchmark data and O is the observed diversity calculated

on the observed sampling data. This index goes to zero when observed diversity = 0, and will rise above 100 if observed diversity is higher than reference diversity.

1.2. Species intactness indices based on occurrence

Occurrence indices measure intactness by comparing the proportion of the sites in a defined region that are occupied by a species between a reference and an observed dataset. Occurrence indices are not intended to track whether a particular site is occupied; rather these indices evaluate the proportion of sites occupied at a landscape scale. All of the occurrence indices are scaled from 0 (degraded) to 100 (intact). The distributions of these indices are shown in Fig. 1.

(4) *Nielsen occurrence index*: Nielsen et al. (2007) proposed an occurrence index where O_i is the observed occurrence rate, R_i the reference (intact) occurrence rate, and s the number

of species. For species observed above the reference rate the index is:

$$\text{Nielsen-OI} = \frac{1}{s} \sum_i [100 - (|O_i - R_i| \times 100)]$$

For species observed below the reference rate the index is:

$$\text{Nielsen-OI} = \frac{1}{s} \sum_i \left\{ 100 - \left[\frac{|O_i - R_i|}{R_i} \times 100 \right] \right\}$$

Two equations are needed for the Nielsen-OI to accommodate both the degradation in biodiversity intactness caused by invasive or weedy species that are present at more sites than normal, and the degradation caused by other species declining in abundance. Nielsen et al. (2007) also combined their occurrence and abundance indices into a "species intactness index" prior to averaging across all species. We, however, have chosen to average across species within each of the occurrence and abundance indices to more fully evaluate the utility of abundance vs. occurrence indices.

(5) *Buckland geometric occurrence index*: Buckland et al. (2005) proposed two indices similar to Nielsen et al. (2007) that can be used with occurrence data where O_i is the observed occurrence rate, R_i the reference (intact) occurrence rate, and s the number of species. The Buckland geometric occurrence index is estimated for species observed below their reference occurrence rate as:

$$\text{BuckGeo-OI} = \left[\exp\left(\frac{1}{s} \sum_i \log\left(\frac{O_i}{R_i} \right) \right) \right] \times 100$$

Buckland's index as published increased above 100 when a species was more prevalent than expected. We modified their index to accommodate species that rose above their reference occurrence rates by reversing the numerator and denominator for those species. This modification is important because invasive and weedy species that increase above their reference conditions are as important an indicator of declining biodiversity intactness as a native species declining below its reference conditions. This modification constrained the index to range between 0 and 100 and makes it directly comparable to the other indices we examined. Buckland's index as published scaled between 0 and 1 so we multiplied the value by 100 to scale it the same as the other indices examined. A practical issue presented by the BuckGeo-OI index is in situations where the observed or reference occurrence is zero. We have chosen to substitute 0.05 for zero in those cases to avoid taking a logarithm of zero.

(6) *Buckland arithmetic occurrence index*: the Buckland arithmetic occurrence index for species observed below their reference occurrence rate is:

$$\text{BuckArith-OI} = \frac{1}{s} \sum_i \left(\frac{O_i}{R_i} \right) \times 100$$

Like the Buckland geometric index, the numerator and denominator in the above relationship are reversed when the observed occurrence is higher than the reference occurrence. The BuckArith-OI index yields the same values as the BuckGeo-OI index for a single species; only when

Fig. 1 – Behavior of the occurrence (A) and abundance (B) species intactness indices over a range of values. Note that for a single species the behavior of the Buckland geometric and Buckland arithmetic indices are identical. The occurrence indices were calculated based on a reference occurrence rate of 0.5, or the rate where the index value equals 100. The abundance indices were calculated based on a reference occurrence rate (intact) of 100, or the rate where the index value equals 100.

averages are taken across multiple species do the indices differ.

1.3. Species intactness indices based on abundance

Abundance indices measure intactness by comparing the average population size at occupied sites between reference and observed datasets. This can be done by including sites where the species does not occur. However, in most biodiversity intactness monitoring programs there is a very high prevalence of sites where species do not occur. This results in an excessive number of zeros and creates analytical problems using standard statistical distributions. An alternative is to use only those sites that are occupied when calculating abundance indices. This reduces the skewness of the data distribution and allows the factors that influence occurrence to be separated from those that influence abundance (Wright, 1991). The calculation of abundance data on only occupied sites is also critical if occurrence and abundance indices are to be combined to calculate an overall species intactness index (Nielsen et al., 2007). All of the abundance indices discussed in this paper are rescaled to range between 100 (intact) and 0 (completely degraded). The distributions of these indices are shown in Fig. 1.

(7) *Nielsen abundance index*: Nielsen et al. (2007) proposed an abundance index that could accommodate both declining and invading species where A indicates observed abundance, E reference abundance, and s the number of species:

$$\text{Nielsen-Abund} = \frac{1}{s}\sum_i \left\{ 100 - \left[\left(\frac{|(A_i+0.5)^{0.5}-(E_i+0.5)^{0.5}|}{(E_i+0.5)^{0.5}}\right)\times 100\right]\right\}$$

(8) *Nielsen maximum abundance index*: the Nielsen-Abund index has the undesirable property that it must be set to zero when the square root of $A_i + 0.5$ is greater than two times the square root of E_i. We therefore suggest a modified index where the larger of A_i and E_i is substituted into the denominator:

$$\text{Nielsen-Max-Abund} = \frac{1}{s}\sum_i \left\{ 100 - \left[\left(\frac{|(A_i+0.5)^{0.5}-(E_i+0.5)^{0.5}|}{(\max(A_i,E_i)+0.5)^{0.5}}\right)\times 100\right]\right\}$$

The Nielsen-Max-Abund index is identical to the Nielsen-Abund index when observed abundances are less than intact. This modification does not correct the problem that neither of these indices naturally reach zero when a species is extirpated. Thus, the indices are set to zero if the reference abundance is greater than zero and the observed is zero (species extirpated). When the reference and observed abundances are both zero both of these indices are set to 100.

(9) *Buckland geometric occurrence index* and (10) *Buckland arithmetic occurrence index*: the occurrence indices proposed by Buckland et al. (2005) can be applied without modification to abundance data. We evaluate both geometric mean

(BuckGeo-Abund) and arithmetic mean (BuckGeo-Arith) abundance indices in this paper.

1.4. Multivariate community intactness indices

The species and biodiversity intactness indices described above assume that changes in biodiversity intactness at larger spatial and taxonomic scales can be effectively summarized by averaging the occurrence or abundance indices of individual species across all species in the community. In this approach information on community patterns such as species co-occurrence is lost. Multivariate indices that compare the deviation of an observed community structure from the corresponding reference community structure incorporate this information. Many methods of multivariate comparison can be used including rank abundance curves, ordination, taxonomic diversity, Mantel tests, the Procrustes statistic, and blocked multiple-response permutation procedures (MRBP) (Clarke, 1990; Clarke and Warwick, 1998; Peres-Neto and Jackson, 2001; McCune and Grace, 2002; Magurran, 2004; Hewitt et al., 2005; Flåten et al., 2007). Here we test two multivariate indices based on the standardized Mantel statistic and principal components analysis (PCA) ordination. We chose the Mantel test statistic and PCA methods for testing because both tests are familiar to most ecologists, and both provide output that is readily converted to indices ranging from 0 to 100.

(11) *Mantel Sorenson index* and (12) *Mantel Bray-Curtis index*: the Mantel index relies on comparisons of similarity or dissimilarity matrices. The Mantel statistic is simply the Pearson correlation between two matrices (Mantel, 1967; McCune and Grace, 2002). The Mantel statistic ranges between 1 and −1, and in practice is generally limited to between 1 and ∼0. A value of 1 indicates that similarity matrices taken from the observed and reference communities are perfectly correlated, while a value of zero indicates no relationship between the matrices. Negative correlations near zero are common when there are no similarities between communities, but more extreme negative correlations are likely to be rare. With a range between 1 and ∼0, the Mantel statistic is readily converted to an index ranging from 100 to 0. Unlike the species intactness indices the Mantel statistic tracks changes in the community at the individual site level rather than estimating an average or proportion across a region.

The choice of the similarity or dissimilarity measure used to calculate the matrices used in the Mantel test are important. A linear relationship between the distance matrices is assumed, and an inappropriate measure can obscure patterns in the raw data (McCune and Grace, 2002). Numerous distance measures are available and advice on the most appropriate ones can be conflicting (e.g. Faith et al., 1987; Legendre and Legendre, 1998; McCune and Grace, 2002; Kenkel, 2006).We initially evaluated four distance measures: Euclidian and Bray-Curtis (Sorenson) distance for abundance data, and Jaccard and Sorenson distance for presence–absence data. In our simulations we found that the biodiversity intactness indices were affected by the type of data (presence–absence or full community datasets) but were largely unaffected by the choice in distance measure so we only present results for Bray-

Curtis distance using abundance data (Mantel-BC) and for Sorenson distance using presence–absence data (Mantel-Sor).

(13) *Principal component index*: the principal components analysis (PCA) ordination-based biodiversity intactness index assesses the deviation in structure between the reference and observed communities by comparing differences in the positions of the centroids of the observed and reference communities on the first PCA axis. A PCA ordination (Manly, 1994) of the reference community (a set of sampling locations with reference abundances for two or more species) is computed without centering or other standardizing options, and the centroid, or average position (score) in ordination space, of all samples on the first principal component is identified. This could be the community as assessed at the first period of monitoring. The centroid of the observed dataset is identified in the reference ordination space by multiplying average abundance of each species in the observed dataset by the eigenvectors from the scores estimated by the reference ordination (Manly, 1994). If there are no differences between the observed and reference data then the observed centroid should be in the same location in multivariate space as the reference centroid. Increasing differences between the observed and reference data will increase the distance between the centroids. We have used the difference in position between the two centroids to create an index ranging between 100 (intact) and 0 (degraded) as follows:

$$PCA = \left(1 - \left\{ \frac{[\max(|C_E|, |C_O|) - \min(|C_E|, |C_O|)]}{\max(|C_E|, |C_O|)} \right\} \right) \times 100$$

C_E is the position of the expected (reference) centroid and C_O is the position of the observed centroid on the first ordination axis. In the simulated datasets used in this study we found that only the first principal component generated meaningful index values. The first component generally had an eigenvalue far larger than the second component, indicating a very simple structure to the data. The simple structure of the datasets generated in this study is a consequence of there being no biological associations between the randomly generated species. In real monitoring data it may be necessary to incorporate centroid positions along more than one axis.

2. Index evaluation

2.1. Data simulator

Using Microsoft Excel we created random datasets of observed and reference data to test the effectiveness of the biodiversity intactness indices. These simulated datasets were designed to be representative (i.e. similar numbers of species, sites, initial occurrence rates, and initial abundances) of the data expected to be collected by the Alberta Biodiversity Monitoring Institute (www.abmi.ca). Simulated datasets were used rather than real data because these simulated data placed all aspects of the data under our control. This provided a consistent framework with known properties to evaluate the relative performance of the indices. Each dataset was created by randomly populating 100 sites with 62 species. Each species was randomly assigned to sites until the number of occupied sites matched its initial occurrence rate. If a species was selected to be present at a site,

its abundance at that site (number of individuals) was determined by selecting a random integer from a Poisson distribution with a mean equal to the initial abundance for that species. Species associations (non-random co-occurrences of certain species) were not introduced into these data resulting in simulated data without a strong community structure. The initial occurrence rates and abundances of each species were preset by the user. This initial dataset became the reference or intact dataset to which ''observed'' datasets were compared. Observed datasets differed from the corresponding reference dataset through species detection error, the invasion or extinction of one or more species, or along a 50-year time series due to increases or declines in species abundance. The PopTools add-in v. 2.7.5 (http://www.cse.csiro.au/poptools) was used to generate random variables and to carry out matrix functions such as distance matrix calculation and PCA.

2.2. Detection error

Detection error, where a species was present at a sampling location but not observed is a frequent source of error in field surveys (Yoccoz et al., 2001). We evaluated the sensitivity of each index by introducing species detection error into the observed dataset. The reference dataset was assumed to be measured without error. Detection error was introduced in two stages. First, the mean detection rate across all species was set. Second, the detectability of individual species, or the proportion of sites where the species was present but not observed, was determined by randomly selecting a value from a normal distribution with a mean equal to the dataset-wide detection rate and a standard deviation of 0.2. Detection rates greater than 1 and less than 0 were set to 1 and 0.

The sensitivity of the indices to detection error were first examined by plotting the relationship between index value and detection error rate in 1000 datasets where the observed dataset differed from the corresponding reference dataset (data without error) only due to detection error. In this test the mean detection rate was set as a random number selected from a uniform distribution ranging between 0 and 100%. Next, we examined the effects of detection error on the power of indices to detect trends (see below). In these tests the mean detection rate was set to be either high (100%), medium (75%), or low (50%). Due to truncation of species detection values at 1 and 0, these tests were done at actual mean detection rates across all species of ∼0.92 (high), ∼0.74 (medium), and ∼0.50 (low).

2.3. Species invasion or extinction and community turnover

It is important that indices of diversity intactness can detect patterns of species invasion, extinction, and community turnover. We compared the performance of the indices under six scenarios (Table 1). In the species invasion scenarios one or more species of the 62 total species were selected to invade (i.e. were present in the observed occurrence and abundance data set while reference occurrence and abundance were set to zero). Similarly, in the extinction scenarios one or more species were randomly selected to be extirpated (observed occurrence and abundance set to zero). In the turnover

#	Scenario	Description	Expected index value
Table 1 – The six ecological scenarios used to evaluate the indices			
1	Control	The only differences between reference and observed are due to detection error	100-Detection error
2	1 extirpation	1 randomly selected species is extirpated (missing from the observed data)	Small decline below control
3	1 Spp. turnover	1 randomly selected species invades (missing from reference data) and 1 other species is extirpated.	Small decline below scenario 2
4	20% invasion	∼20% of the 62 species randomly selected to invade	Strong decline below control
5	20% turnover	∼20% of the 62 species randomly selected for extirpation and another ∼20% selected for invasion	Moderate decline below scenario 4
6	Complete turnover	∼50% of the 62 species randomly selected for extirpation and the remaining species selected for invasion	Very large decline; may reach zero

The expected differences in the index values for scenarios 2 through 6 are qualitative. See the section below on how low an index should go for a discussion of whether the expected value for scenario 6 should be zero.

scenario one or more species were randomly selected for extirpation while a similar number of other species were selected to invade. In the most extreme scenario complete turnover of the community was simulated by randomly selecting ∼50% of the species for extirpation and setting all of the remaining species to be invaders. In all of these cases the overall species pool remained the same (62 species), but in the invasion and turnover scenarios the starting community contained fewer species. In each case index values were calculated comparing the observed and corresponding reference datasets. Detection rates were set to be high (∼0.92) and the indices were calculated for 500 separate simulated communities. Index performance in these scenarios was evaluated as the mean and variation in index values over the 500 model runs.

2.4. Population trends

We examined the power of the indices to detect trends in intactness by creating 50-year time series with known rates of population change. The dataset-wide mean rate of population change (% change in abundance per year) across all species was first set in the model. Rates of change in occurrence were not directly modeled; rather changes in abundance led to sites become occupied or unoccupied. The population trends for individual species were set by selecting random numbers from a normal distribution with a mean equal to the dataset-wide mean rate of abundance change and a standard deviation of 1. A second random number determined whether each species would increase or decrease in the population. These trends remained fixed throughout the entire time series. Approximately 50% of the species in any time series were increasers and approximately 50% decreasers. In each year of the time series the abundance of each species at each sample location changed from the previous step following a random number drawn from a Poisson distribution with a mean equal to the overall rate of change for that species. Decreasing species could decline to extinction, and once extinct it was not possible for that species to return.

We ran 500 time series at two dataset-wide mean trend rates (1 and 3% change in abundance per time step) and three

levels of detection error (∼92, ∼74, and ∼50% detection rates). We used the STATSBY command in STATA Version 9 (StataCorp; College Station, TX) to estimate a linear regression model for each community index predicted by time step for the 500 simulations. Regression models were run for 3 different lengths of time: 5 time steps, 25 time steps, and 50 time steps. We recorded the number of times linear models for community indices exceeded the critical F-value based on statistical significance of 0.95 ($\alpha \leq 0.05$). If this criteria was met in at least 90% of the 500 simulations we considered the index to have high statistical power to detect the time trend ($\beta = 0.1$).

3. Results

3.1. Detection error

All of the indices were affected by detection error (Fig. 2). The species richness, Nielsen-OI and BuckArith-OI indices had nearly linear relationships between index value and average detectability. The BuckGeo-OI index, however, declined more precipitously relative to the other dices as detection error increased. Both the Shannon and Simpson diversity indices and the abundance indices were generally much less sensitive to detection error than the occurrence indices. The Nielsen-Abund and BuckArith-Abund were only strongly influenced when detectability dropped below 0.5. The BuckGeo-Abund index was consistently more sensitive to detection error than the other abundance indices. The abundance indices were less sensitive to detection error than occurrence indices. Only when detection error was very high and mean abundance quite low did false zeros strongly impact the observed average abundance for a species. The PCA index was generally less sensitive to detection error than the Mantel-based indices.

3.2. Species invasion or extinction and community turnover

The BuckGeo-OI and BuckGeo-Abund indices effectively portrayed all of the declines in intactness that were simulated
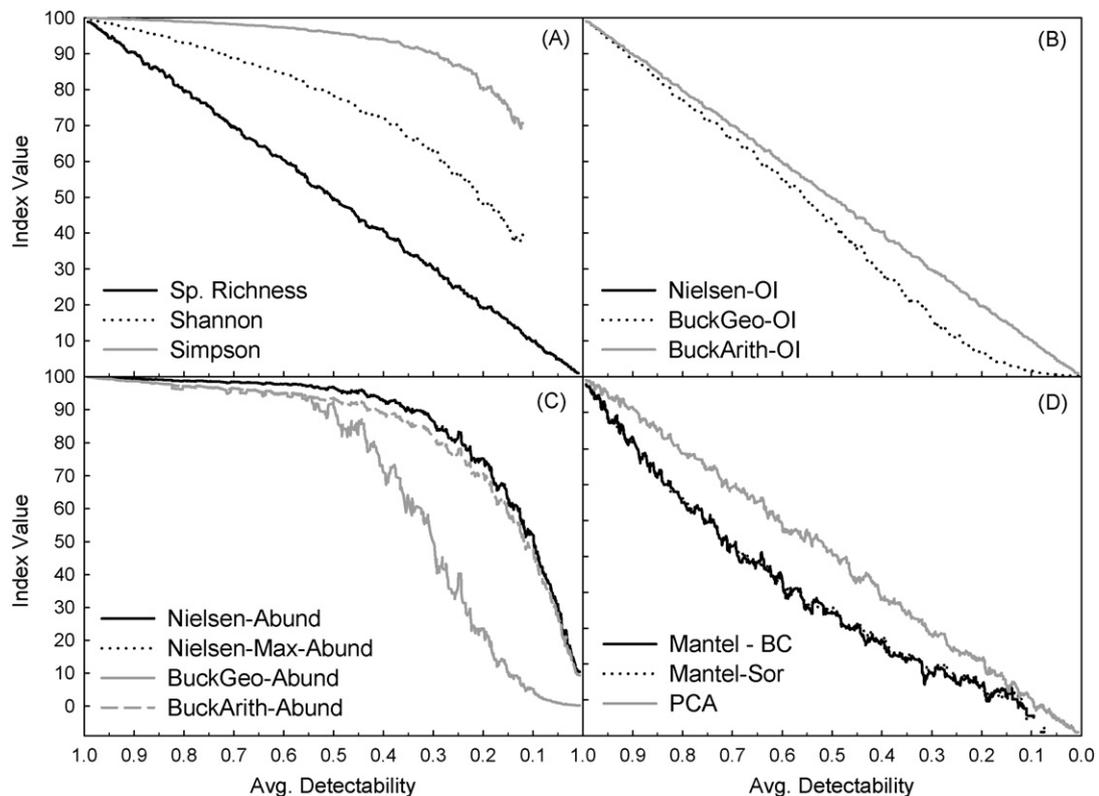
Fig. 2 – Performance of the diversity indices (A), occurrence indices (B), abundance indices (C), and multivariate indices (D) across a range of detectability (1000 model runs with random levels of detection error). The lines in the figure follow the moving average (window of 10 observations) for each index with increasing detection error. The Shannon and Simpson's diversity indices were not calculated at very low detectability due to the presence of sites with no detected species. The performance of the Nielsen-OI and BuckArith-OI indices were identical, and the Nielsen-Abund and Nielsen-Max-Abund were similarly nearly identical.

(Fig. 3). The traditional indices of diversity were ineffective in all scenarios, while the other species intactness and community intactness indices were only effective at 20% or more invasion or turnover. Two indices failed to reach zero in scenario six where all native species were replaced by invasive species (complete turnover). The Nielsen-OI index only declined to 26.30 ± 3.59 (S.D.) and the Nielsen-Max-Abund index declined to 16.55 ± 2.02 (S.D.). The Mantel-based community structure indices both performed similarly across the six scenarios, and reached average values very close to zero with complete turnover: Mantel-BC (abundance) 0.21 ± 4.17 (S.D.); Mantel-Sor (presence–absence) -0.02 ± 4.08 (S.D.). The PCA index was less sensitive than the Mantel-based indices in some scenarios, but reached zero with complete turnover.

3.3. Species trends

On average, the power of the indices to detect the negative trends in intactness that were simulated increased with the length of the time series and declined as detection error rates increased (Tables 2 and 3 and Fig. 4). The species abundance indices had consistently high power to detect trends at all trend rates and levels of detection error (Table 4). The species occurrence indices also had high power to detect trends when detection error was low, but power was strongly reduced by high levels of detection error. All of the species occurrence indices performed similarly, though the BuckArith-OI had higher power with low rates of detection error while the Nielsen-OI had higher power at high rates of detection error. Traditional indices were generally ineffective for detecting trends in intactness as the abundance of weedy species lead to increases in species richness and diversity. Finally, the Mantel-based multivariate indices were effective for detecting trends but the PCA-based multivariate index performed very poorly. In particular the PCA gave misleading results at high levels of detection error (Fig. 4).

Both the occurrence indices and the PCA index had hump-shaped trends in intactness at higher levels of detection error. Intactness in these indices appeared to increase early in the time series (Fig. 4) resulting in apparently higher power to detect changes in short (5 year) time series than in longer series (Tables 2 and 3). The apparent increases in intactness were caused by population trends in the species with occurrence rates that had risen above their reference occurrence rates. These species were actually declining in intactness, as shown by the trends at high detection rates, because their true occurrence rates had moved above their reference values. At lower rates of detectability however, the initially observed occurrence
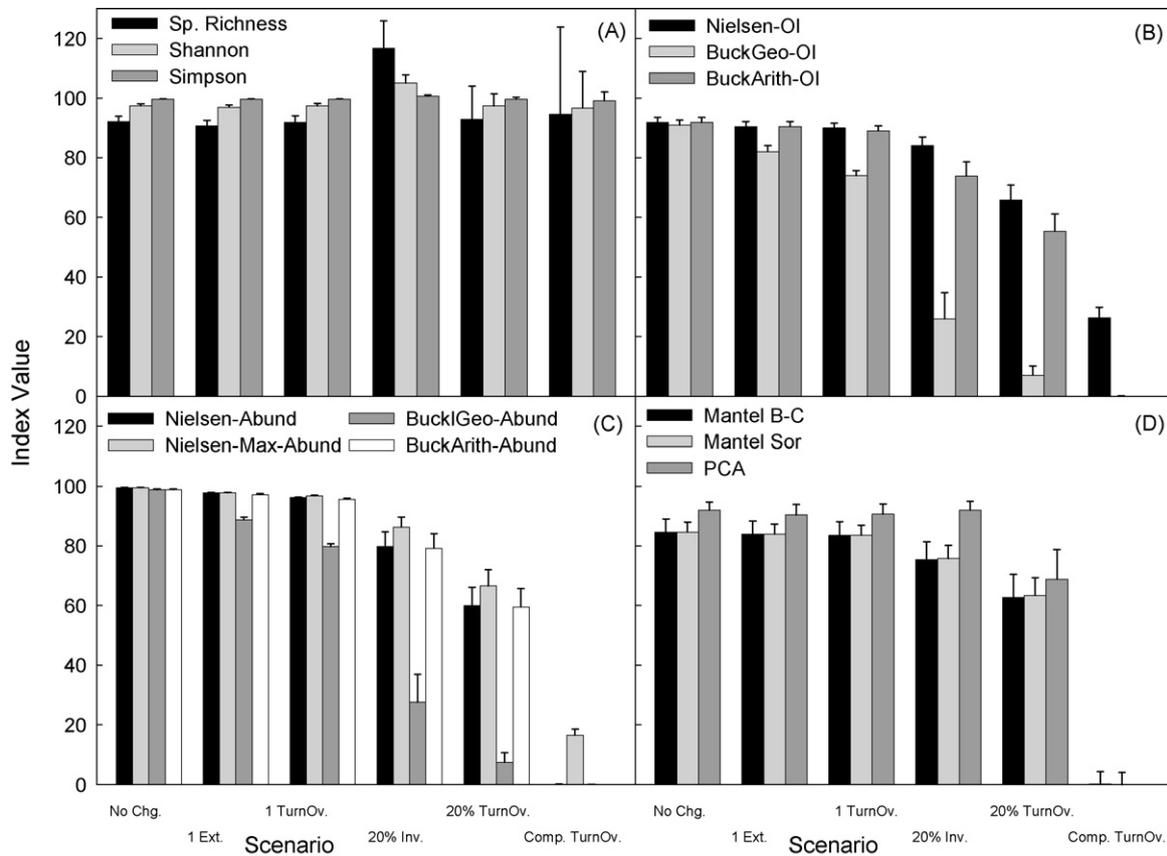
Fig. 3 – Performance of the diversity indices (A) occurrence indices (B), abundance indices (C), and multivariate indices (D) in each of six ecological scenarios (control, 1 species extirpated, turnover of 1 species, 20% of the species in the community invasive, 20% turnover, and complete turnover). Bars are the mean index value over 500 model runs and the error bars are 1 standard deviation.

rates were reduced below intact values because many populations were not detected. Intactness appeared to rise early in the time series because the increased real abundances led to more detected populations, and hence occurrence rates similar to the reference rates. A decline in intactness was only observed when the species became so abundant that, even with detection error, the observed occurrence rate rose above the reference rate. The abundance indices were not affected because those indices are calculated from the average abundance values from only the observed sites.

4. Discussion

4.1. Choice of index

We ranked the 13 indices assessed in this paper based on statistical power, ecological relevance, and ease of communication (Table 4). These rankings suggest that the species abundance indices are the best overall, followed by the species occurrence indices. Within each category the best performing indices were species richness (traditional diversity indices), Buck-Arith-OI (species occurrence indices) Buck-Arith-Abund

(species abundance indices) and Mantel-Sor (multivariate indices).

Species richness and the Shannon and Simpson indices are generally not useful for large-scale monitoring of diversity intactness. Since information on species identity is lost, they cannot be used to monitor species turnover. Species richness ranked highly based on its power to detect a significant trend (Tables 2 and 3), however the trend detected was an "improvement" due to the arrival of non-native species (Fig. 4). These problems can be solved on an ad hoc basis by creating separate indices for native and non-native species, but given that more effective indices are available there is little reason to take this approach. The simplicity of species richness can enhance communication in some settings, but in a monitoring program these indices are likely to pose barriers to effective communication as they are not readily converted to a common (0–100) scale for comparison. By the time that richness and diversity indices detect any change in the state of biodiversity it is likely to be too late to do anything about it.

The BuckArith-OI and BuckArith-Abund indices were ranked as the best occurrence and abundance indices, respectively. The BuckGeo-OI and BuckGeo-Abund indices

Table 2 – Power results for biodiversity indices along time series (TS) of three lengths (5, 25, and 50 years) with a −1% trend in intactness per year

	Detection High			Detection Medium			Detection Low		
	TS5	TS25	TS50	TS5	TS25	TS50	TS5	TS25	TS50
Traditional diversity indices									
Sp. Rich	0.98	1	1	0.798	1	0.996	0.528	0.996	0.870
Shannon	0.800	0.966	0.814	0.382	0.968	0.788	0.288	0.804	0.822
Simpson	0.466	0.832	0.906	0.178	0.758	0.880	0.122	0.692	0.914
Species Occurrence indices									
Nielsen-OI	0.598	1	1	0.173	0.952	1	0.390	0.788	0.586
BuckArith-OI	0.938	1	1	0.098	0.994	1	0.286	0.448	0.868
BuckGeo-OI	0.886	1	1	0.060	0.996	1	0.162	0.106	0.878
Species Abundance indices									
Nielsen-Abund	0.998	1	1	0.856	1	1	0.286	1	1
Nielsen-Max-Abund	0.998	1	1	0.862	1	1	0.300	1	1
BuckArith-Abund	1	1	1	0.860	1	1	0.360	1	1
BuckGeo-Abund	1	1	1	0.836	1	1	0.212	0.792	0.998
Multivariate Community indices									
Mantel-BC	0.194	1	1	0.078	1	1	0.052	0.862	1
Mantel-Sor	0.956	1	1	0.444	1	1	0.144	1	1
PCA	0.264	0.822	0.900	0.114	0.896	0.940	0.068	0.774	0.958

Power is measured as the (proportion of simulation runs [n = 500] with P-value <0.05 at three detection rates: high (∼0.92), medium (∼0.75) and low (∼0.50). Shaded cells indicate a power value of 0.9 or greater.

were more sensitive to detection error, and they require that an arbitrary value be entered when one or both of the observed or reference values are zero to avoid taking the logarithm of zero. The Nielsen-Abund index has the undesirable property that to avoid negative index values the index must be set to zero when the square root of observed species abundance is greater than two times the square root of expected abundance. Both the Nielsen-OI and Nielsen-Max-Abund indices fail go to zero when a native community is completely replaced by non-native species. While it is not a necessity that indices should go to zero in this situation, as discussed in more detail below, it is likely to ease communication if they do reach zero.

We found that the Mantel-Sor multivariate index had higher power to detect trends than the PCA and Mantel-BC indices, and was not affected by time-detection error interactions. The severe impact of time-detection error interactions on the PCA index makes it only practical for intactness monitoring in situations where users have a very high degree of confidence that detectability is close to 1.

4.2. Management of detection error

All of the tests in this manuscript assumed a perfectly measured reference dataset and an observed dataset that contained a degree of detection error. Substantial improvements in statistical power could likely be achieved by incorporating detection error into the reference dataset and estimating detection error in any analysis of trends. These increases in power would arise because the rates of site occupation and average abundance in the reference dataset would be likely adjusted by missed observations occurring at similar rates to those in the observed data. Examples of reference data where detection error would occur include cases where the reference dataset is the first observation on a time series or an undisturbed location (e.g. Sinclair et al., 2002; Loh et al., 2005; Scholes and Biggs, 2005). Similarly, in cases where empirically derived reference datasets are used, detection error can be explicitly incorporated into the reference dataset (Nielsen et al., 2007). Finally, it is important to note that indices with high sensitivity to detection error are also very sensitive to small changes in diversity intactness. In cases where detection error is low or remains constant over a series of surveys, then selection of a more sensitive index may improve statistical power. Future research needs to investigate the improvement in statistical power that occurs in biodiversity trends by incorporating equations that correct for detection error (e.g. MacKenzie et al., 2005).

Table 3 – Power results for biodiversity indices along time series (TS) of three lengths (5, 25, and 50 years) with a −3% trend in intactness per year

	Detection High			Detection Medium			Detection Low		
	TS5	TS25	TS50	TS5	TS25	TS50	TS5	TS25	TS50
Traditional diversity indices									
Sp. Rich	1	0.988	0.630	0.984	0.992	0.654	0.858	0.968	0.658
Shannon	0.952	0.680	0.944	0.798	0.658	0.906	0.464	0.650	0.872
Simpson	0.640	0.846	1	0.364	0.830	0.994	0.194	0.780	0.998
Species Occurrence indices									
Nielsen-OI	0.988	1	1	0.132	1	1	0.472	0.794	1
BuckArith-OI	1	1	1	0.566	1	1	0.258	0.988	1
BuckGeo-OI	1	1	1	0.456	1	1	0.118	0.978	1
Species Abundance indices									
Nielsen-Abund	1	1	1	0.980	1	1	0.546	1	1
Nielsen-Max-Abund	1	1	1	0.982	1	1	0.548	1	1
BuckArith-Abund	1	1	1	0.996	1	1	0.716	1	1
BuckGeo-Abund	1	1	1	0.966	1	1	0.472	0.996	1
Multivariate Community indices									
Mantel-BC	0.646	1	1	0.214	1	1	0.096	1	1
Mantel-Sor	0.992	1	1	0.780	1	1	0.346	1	1
PCA	0.480	0.848	0.982	0.284	0.890	0.840	0.136	0.952	0.930

Power is measured as the (proportion of simulation runs [$n = 500$] with P-value <0.05 at three detection rates: high (∼0.92), medium (∼0.75) and low (∼0.50). Shaded cells indicate a power value of 0.9 or greater.

4.3. Multivariate vs. univariate indices of diversity intactness

There are many theoretical reasons to prefer multivariate indices of diversity intactness over the univariate indices discussed in this paper. The structure of communities and the interactions among species are as much a feature of biodiversity as species lists, and changes in community structure can alter the provision of many important ecosystem services (Chapin et al., 2000). While multivariate indices of biodiversity intactness may be preferred for theoretical reasons, the higher sensitivity of multivariate indices, particularly the PCA index, to detection error makes their practical application challenging. The species intactness approaches used in this paper can be "corrected" for detection error through methods like multiple visit sampling (MacKenzie et al., 2002). No such corrections exist for multivariate indices to our knowledge. In addition, the mathematical complexity of the multivariate indices may pose barriers to effective communication. In particular some users may feel that the complex procedures were used to obscure trends. This is not a risk with the indices based on simple arithmetic operations. Multivariate indices also need to be calculated anew at every level of spatial and taxonomic resolution, making cross-scale comparisons more difficult.

4.4. Monitoring of individual sites vs. monitoring of landscapes

The indices presented in this paper are designed to monitor biodiversity intactness at the landscape scale. This is a departure from most monitoring programs, which are focused on the changes in diversity occurring at intensely studied individual sites. Using indices based on occurrence rates and abundances averaged across a large number of sites has a number of advantages. First, detection error becomes less problematic when large numbers of sites are monitored. If only a few sites are monitored then small errors in detection can result in substantial changes in trends. With large-scale monitoring programs with many sites, the effects of detection error are averaged across sites, reducing the influence of detection error at any one point and making the trends more robust. This approach also allows a shift in sampling strategy from the intensive monitoring of a small number of sites to less intensive monitoring of many sites (i.e. assessing the state of a large statistical population). Monitoring many sites also better reflects the ecology of many species. In a healthy population of almost any species there are likely to be many apparently suitable sites that are not occupied (e.g. Hanski, 1998; Freckleton and Watkinson, 2002). The observation that an individual site may be occupied on one survey but not on another provides little information on the intactness of that
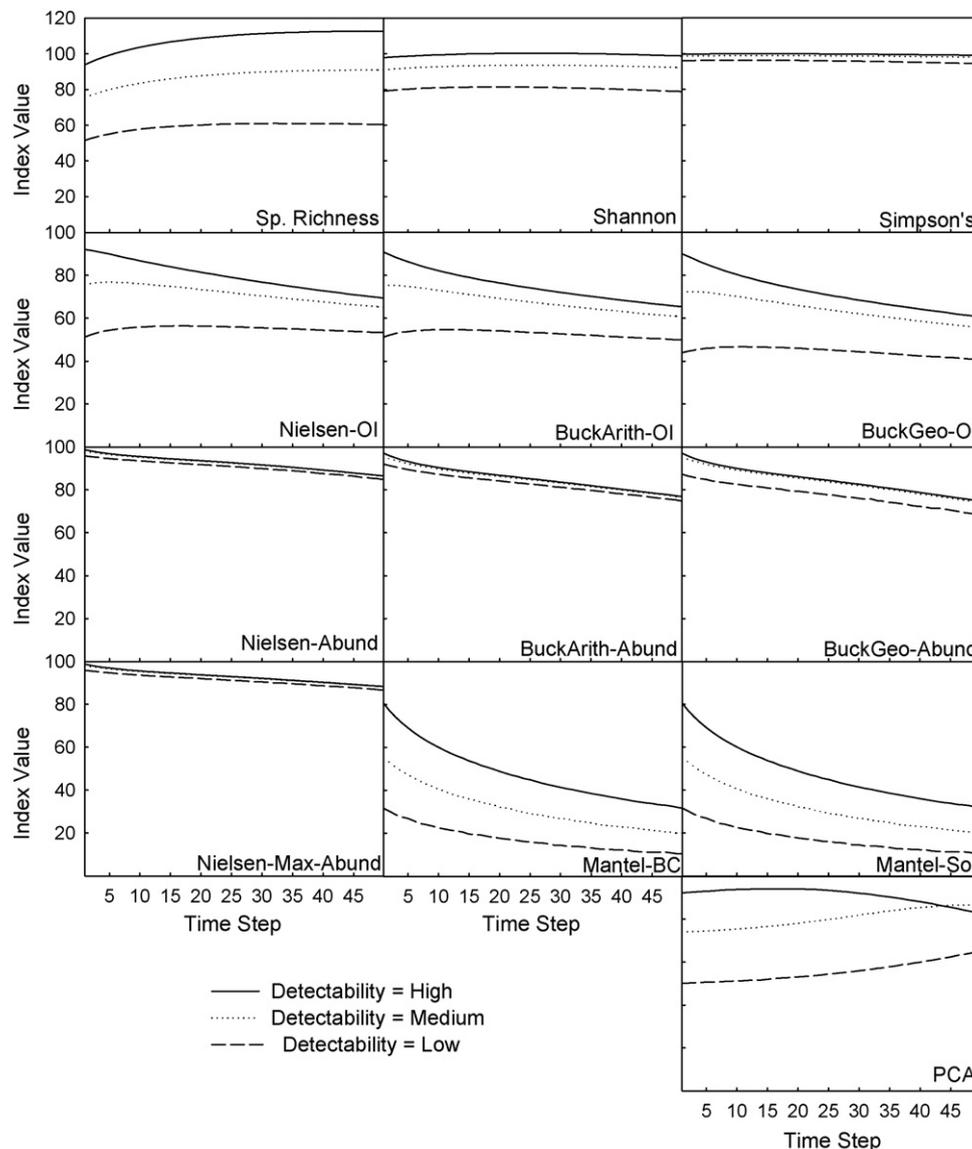
Fig. 4 – Performance of the diversity, occurrence, abundance, and multivariate indices in a 50 step time series with intactness declining by 1% per step. Values for each time step are an average of 500 time series. The performance of each index is shown for high, medium, and low rates of detectability.

species. A site may be unoccupied because it has been degraded and hence no longer is suitable habitat, or it may be unoccupied because the previous occupant had died or moved on, but a new occupant had not yet dispersed to the site. With an appropriate sampling design, we gain much more information about the intactness of a species if we understand trends in occurrence rates and abundance at a landscape scale than small-scale monitoring done with higher temporal resolution.

4.5. How low should an index go?

Several of the intactness indices discussed in this paper reach zero when all of the native species in a community are replaced by non-native species (species with a reference occurrence and abundance of zero) while others indices do not

(Fig. 4). The issue of whether complete turnover should result in a biodiversity intactness value of zero is important because, while the extirpation of all native species is clearly a severe degradation of biodiversity intactness, species are still present and providing ecosystem services. Intactness indices that go to zero treat a non-native community as equivalent to degradation so severe that no species at all are present. While this argument suggests that a biodiversity intactness index should not go to zero with complete turnover, it is not clear to what value the index should fall to. We suggest using biodiversity intactness indices that fall to zero with complete turnover. Intact biodiversity is defined as the community found in the absence of anthropogenic disturbance (Nielsen et al., 2007). By this definition a community composed entirely of non-native species is highly degraded and should have an intactness value close to zero. In addition, a diversity

Table 4 – Summary of index performance

Index	Detection error[a]	Power (−1% trend)[b]	Power (−3% trend)b	Ecological scenarios[c]	Arbitrariness[d]	Communication[e]	Overall
Traditional diversity indices							
Sp. rich	8.5	5	10	11.5	5	3.5	7
Shannon	3.5	12.5	12	11.5	5	8.5	11
Simpson	5.5	12.5	12	13	5	8.5	12
Species occurrence indices							
Nielsen-OI	8.5	9.5	8.5	9	5	8.5	9
BuckArith-OI	8.5	7.5	6	5	5	3.5	5
BuckGeo-OI	11	9.5	6	1.5	11.5	3.5	6
Species abundance indices							
Nielsen-Abund	1.5	2	2.5	5	11.5	3.5	2
Nielsen-Max-Abund	1.5	2	2.5	9	5	8.5	3
BuckArith-Abund	3.5	2	2.5	5	5	3.5	1
BuckGeo-Abund	5.5	5	2.5	1.5	11.5	3.5	4
Multivariate community indices							
Mantel-BC	12.5	7.5	8.5	5	5	12	10
Mantel-Sor	12.5	5	6	5	5	12	8
PCA	8.5	11	12	9	11.5	12	13

In each category indices were ranked relative to one another from 1 (best) to 13 (worst). Note that all ranks were corrected for ties.
[a] Ranks based on the shape of the relationships in Fig. 2. Convex curves were ranked higher than linear relationships and concave curves were ranked lowest.
[b] Ranks based on the number of tests out of 9 that achieved power ≥0.9.
[c] Ranks based on the number of scenarios that conformed to the expectations described in Table 1.
[d] Ranks based on whether or not the index required an arbitrary mathematical decision.
[e] Subjective ranking based on estimated ease of communication of the index to non-scientists.

intactness index that falls to zero is likely to facilitate communication. Zero intactness defined as the complete absence of native species is a straightforward concept to communicate to users while explaining why the minimum value of an index is greater than zero is likely to be challenging.

5. Conclusions

Statistically rigorous methods to rapidly communicate results to the scientific community, government, industrial stakeholders, and the general public are a critical element of successful biodiversity intactness monitoring programs (Noss, 1990; Debinski and Humphrey, 1997; Yoccoz et al., 2001; Nichols and Williams, 2006; Lovett et al., 2007). In this paper we compare 13 indices based on statistical power, ease of communication, and sensitivity to errors in the data. Our results suggest that intactness index equations based on Buckland et al.'s (2005) arithmetic mean index are the most effective for use in large-scale biodiversity monitoring programs.

REFERENCES

Buckland, S.T., Magurran, A.E., Green, R.E., Fewster, R.M., 2005. Monitoring change in biodiversity through composite indices. Philosophical Transactions of the Royal Society B 360, 243–254.

Chapin III, F.S., Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds, H.L., Hooper, D.U., Lavorel, S., Sala, O.E., Hobbie, S.E., Mack, M.C., Diaz, S., 2000. Consequences of changing biodiversity. Nature 405, 234–242.

Clarke, K.R., Warwick, R.M., 1998. A taxonomic distinctness index and its statistical properties. Journal of Applied Ecology 35, 523–531.

Clarke, K.R., 1990. Comparisons of dominance curves. Journal of Experimental Marine Biology and Ecology 138, 143–157.

Debinski, D.M., Humphrey, P.S., 1997. An integrated approach to biological diversity assessment. Natural Areas Journal 17, 355–365.

Faith, D.P., Minchin, P.R., Belbin, L., 1987. Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69, 57–68.

Flåten, G.R., Botnen, H., Grung, B., Kvalheim, O.M., 2007. Quantifying disturbances in benthic communities—comparison of the community disturbance index (CDI) to other multivariate methods. Ecological Indicators 7, 254–276.

Freckleton, R.P., Watkinson, A.R., 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. Journal of Ecology 90, 419–434.

Hanski, I., 1998. Metapopulation dynamics. Nature 396, 41–49.

Hewitt, J.E., Anderson, M.J., Thrush, S.F., 2005. Assessing and monitoring ecological community health in marine systems. Ecological Applications 15, 942–953.

Kenkel, N.C., 2006. On selecting an appropriate multivariate analysis. Canadian Journal of Plant Science 86, 663–676.

Krebs, C.J., 1989. Ecological Methodology. Harper Collins, USA.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier, Amsterdam.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., Randers, J., 2005. The living planet index: using species

population time series to track trends in biodiversity. Philosophical Transactions of the Royal Society B 360, 289–295.

Lovett, G.M., Burns, D.A., Driscoll, C.T., Jenkins, J.C., Mitchell, M.J., Rustad, L., Shanley, J.B., Likens, G.E., Haeuber, R., 2007. Who needs environmental monitoring? Frontiers in Ecology and the Environment 5, 253–260.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83, 2248–2255.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Hines, J.E., Bailey, L.L., 2005. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence. Elsevier, San Diego, USA.

Magurran, A.E., 2004. Measuring Biological Diversity. Blackwell, Oxford.

Manly, B.F.J., 1994. Multivariate Statistical Methods: A Primer. Chapman and Hall, London.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27, 209–220.

Margalef, R.D., 1958. Information theory in ecology. General Systems 36–71.

McCune, B., Grace, J.B., 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon.

Nichols, J.D., Williams, B.K., 2006. Monitoring for conservation. Trends in Ecology and Evolution 21, 668–673.

Nielsen, S.E., Bayne, E.M., Schieck, J., Herbers, J., Boutin, S., 2007. A new method to estimate species and biodiversity intactness using empirically derived reference conditions. Biological Conservation 137, 403–414.

Noss, R.F., 1990. Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology 4, 355–364.

Overton, J.M., Stephens, R.T.T., Leathwick, J.R., Lehmann, A., 2002. Information pyramids for informed biodiversity conservation. Biodiversity and Conservation 11, 2093–2116.

Peres-Neto, P.R., Jackson, D.A., 2001. How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. Oecologia 129, 169–178.

Scholes, R.J., Biggs, R., 2005. A biodiversity intactness index. Nature 434, 45–49.

Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.

Sinclair, A.R.E., Mduma, S.A.R., Arcese, P., 2002. Protected areas as biodiversity benchmarks for human impact: agriculture and the Serengeti avifauna. Proceedings of the Royal Society of London Series B 269, 2401–2405.

Taft, J.B., Hauser, C., Robertson, K.R., 2006. Estimating floristic integrity in tallgrass prairie. Biological Conservation 131, 42–51.

Wright, D.H., 1991. Correlations between incidence and abundance are expected by chance. Journal of Biogeography 18, 463–466.

Yoccoz, N.G., Nichols, J.D., Boulinier, T., 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16, 446–453.