A unified analysis of niche overlap incorporating data of different types

Shane W. Geange[1]*, Shirley Pledger[2], Kevin C. Burns[1] and Jeffrey S. Shima[1]

[1]*School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand; and*
[2]*School of Mathematics, Statistics and Operations Research, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand*

Summary

1. Competition theory predicts that community structure may be shaped by resource partitioning between co-occurring species. As such, quantifying the degree of resource partitioning (i.e., niche overlap) is a key component of studies examining community structure and species coexistence.

2. For many organisms, multiple resource axes quantify niche space. Each axis may be described by a different type of data (e.g. categorical, continuous, count or binary data, as well as electivity scores), with different data types requiring different statistical treatments. Therefore, incorporating multiple axes into a single measure of niche space is problematic.

3. Here, we propose general methods for combining multiple niche axes, each characterized by different data types, within a unified analysis of niche overlap. Using appropriate transformations and probability models, we show that each data type can give rise to directly comparable measures of niche overlap, with the overlap statistic between two species defined as the overlapping area between the distributions for each species.

4. Measurements derived from different types of data can be combined into a single unified analysis of niche overlap by averaging over multiple axes.

5. We then describe null model permutation tests that assess statistical differences in niche overlap, which can address questions commonly posed by population ecologists (e.g. do two species occupy different niche space?) and community ecologists (e.g. are multiple species evenly distributed across niche space?).

6. To illustrate the use of these newly devised indices, we use an example from reef fishes that combines ratio, categorical and electivity data, and an example from alpine plants that combines continuous and ratio data.

7. The methods described in this article are relevant to a wide variety of ecological projects, including the investigation of invasive species, relative abundance distributions, global change, species coexistence and evolutionary diversification.

Keywords: community structure, multivariate analysis, null model analysis, reef fish, resource use

Introduction

Hutchinson (1957) defined the niche as a multidimensional hypervolume in which a species maintains a viable population. The entire hypervolume under which an organism can potentially exist describes its fundamental niche, whereas the portion of the fundamental niche that a species actually occupies (for example, due to competitive exclusion) defines its realized niche. The breadth of a species' realized or fundamental niche

*Correspondence author. E-mail: shane.geange@vuw.ac.nz
Correspondence site: http://www.respond2articles.com/MEE/

can be viewed in terms of either the Eltonian niche (the functional attributes of a species and its corresponding trophic position: Elton 1927), or the Grinnellian niche (the response of a species to the abiotic and biotic environment: Grinnell 1917). Further, niche overlap can be analysed at a hierarchy of spatial scales: the broad geographical range of a species (the γ niche); the region of a species' niche that corresponds to the habitat(s) where it is found (the β niche); and the region of a species' realized niche corresponding to the local scale where interactions occur among species (the α niche: Silvertown *et al.* 2006). In the case of the local scale, the realized niche of a species may differ between locales where either the identity or abundance

of heterospecifics differs. The related concept of niche overlap between species may be viewed as the volume in multidimensional hyperspace within which two or more species maintain viable populations in the presence of one another (Mouillot *et al.* 2005). The quantification of niche overlap has become an important tool for investigating invasive species (Olden, Poff & Bestgen 2006; Gregory & Macdonald 2009), relative abundance distributions (Sugihara *et al.* 2003), global change (Broennimann *et al.* 2007), species coexistence (Mookerji, Weng & Mazumder 2004; Silvertown 2004) and evolutionary diversification (Day & Young 2004; Smith *et al.* 2004).

Descriptions of niche space often incorporate multiple axes, each of which may be an environmental condition (e.g. altitude, pH); a type of resource (e.g. prey type, refuge type); a phenotypic trait indicating the type of resource used (e.g. gut length may be indicative of diet); or an index of electivity (e.g. Manly's Alpha). Multiple axes may be described by different data types, including: binary data (e.g. presence/absence of a prey type in the diet), categorical data (e.g. host size classes: small, medium or large), continuous data (e.g. soil pH), ratio data (e.g. pectoral fin aspect ratio), count data (e.g. number of prey eaten per hour), or indices of electivity (e.g. habitat use relative to availability). For example, reef fish ecologists may wish to quantify niche space across multiple axes, incorporating swimming ability (pectoral fin aspect ratio bounded below by 0), habitat association (electivity data) and spatial distribution (categorical data). Incorporating multiple axes described by different data types into a single measure of niche space is statistically challenging, because different data types cannot be dealt with in the same way. Here, we propose general statistical methods for combining different data types within a unified analysis of niche overlap, with an emphasis on local realized niches.

Historically, niche overlap has been calculated across single- or multiple axes using either: (1) continuous data (e.g. MacArthur & Levins 1967; Green 1971; Harner & Whitmore 1977; Dueser & Shuggart 1979; Litvak & Hansell 1990); or (2) discrete categorical data (e.g. Pianka 1973; Pappas & Stoermer 1997; Dolédec, Chessel & Gimaret-Carpentier 2000); however, there is currently no methodology for incorporating different data types into a single analysis of niche overlap across multiple niche axes. Continuous data has often been modelled as the overlap of two normal curves (e.g. MacArthur & Levins 1967; Cody 1975; Harner & Whitmore 1977; Manly 1994), although this overlap index has two major problems: (1) it assumes that quantitative functional traits are normally distributed; and (2) it does not account for differences in variance between two normal distributions with the same mean. Although continuous data can be converted to categorical data to overcome these problems, there is a loss of information in replacing measurements with ordered categories. Mouillot *et al.* (2005) provides a solution to these problems, describing an approach, based on kernel distribution estimators, that models niche overlap from continuous data independently of the underlying distribution of the data. Their approach facilitates the construction of broad, multivariate indices of niche overlap. Mouillot *et al.* (2005) constructed their indices from probability distributions of continuous measurements using density estimation to provide comparable metrics across different axes. However, this approach does not provide a framework for other types of data (e.g. electivity scores between 0 and 1, or categorical data), which are often recorded by field workers, and might usefully be included in multivariate indices of niche overlap.

Slobodchikoff & Schulz (1980) also suggest a mathematical basis for calculating overlap across several niche dimensions, which can incorporate both discrete and continuous data. However, their proposed approach was not fully developed. Here, we expand the ideas of Slobodchikoff & Schulz (1980) to develop an approach that provides directly comparable measures of niche overlap from a wide range of data types (e.g. binary, categorical, continuous, ratio, proportion, percent, count and electivity data). Using appropriate transformations and density estimation techniques, each data type gives rise to equivalent measures of niche overlap ranging from 0 (no overlap) to 1 (complete overlap). The use of directly comparable measures of niche overlap ensures that the geometric interpretation of the overlapping density functions or probability is the same for each data type. Once estimated probability distributions are available for each data type, the overlap statistic between two species is simply the overlapping area between the distributions for each species. It is then possible to create a composite output of niche overlap, derived from multivariate inputs of different types of data, by averaging over multiple axes.

A major objective of studies examining niche overlap is to assess differential use of niche space by multiple species. For example, population ecologists may want to know if two species occupy different niche space. Alternatively, community ecologists may want to know if multiple species are evenly spaced across, or clustered within niche space. Niche comparisons must be done in a way that rules out detecting as different two niches which only differ due to sampling variation. We suggest the use of null models (see Gotelli & Graves 1996; Gotelli 2000) to differentiate between species occupying similar and different niches. This approach uses permutation tests to produce a statistical null distribution (the distribution of the test statistic under the null hypothesis of no niche difference) by calculating pseudo-values of the test statistic which would arise if H_0 was true. The position of an empirically derived test statistic in relation to the pseudo-values generated by the null model provides the P-value for the test.

We begin by outlining methodology for combining binary, categorical, continuous, ratio, proportion, percent, count and electivity data into a unified analysis of niche overlap (also see Pledger & Geange 2009). Appropriate transformations and probability models for each data type give rise to estimated probability distributions, with the overlapping area between the probability distributions of two species being the niche overlap statistic. We then outline how individual axes of niche overlap can be combined into a single multivariate measure of niche overlap. We illustrate the construction of composite measures of overlap in local realized niche space (i.e. the α niche) with an example from reef fishes that combines measurement, categorical and electivity data, and an example from alpine plants that combines continuous and ratio data. We use

permutation tests to evaluate whether local realized niche space differs between pairs of species, and if the local realized niches of all species are evenly distributed across niche space.

Materials and methods

DATA TRANSFORMATIONS AND PROBABILITY MODELS

To provide comparable measures of niche overlap over multiple niche axes that incorporate different types of data, we use a combination of mixture models and kernel density estimation. Mixture models, based on a probability distribution (e.g. Bernoulli for binary data or Poisson for count data), provide flexible approximations of the observed distribution for discrete data. Similarly, kernel density estimation, an extreme case of finite mixtures when the number of components in the mixture is large, is used to approximate the observed distribution for continuous data. The advantage of using these methods is that they are amenable to using the same formula for niche overlap; therefore, they can be used to provide directly comparable measures of niche overlap. Here, equations 1–4 use either mixture models or kernel density estimation to provide directly comparable measures of niche overlap across discrete and continuous data, ranging between 0 and 1. Of note is that: equation 1 is a special case of equation 2; and equations 2 and 4 are essentially the same (they differ only in how they are used to model discrete and continuous data, respectively).

Binary data

Many measures of resource overlap are binary. An example is a habitat patch that may be colonized or not (e.g. Moilanen & Nieminen 2002). For such data, we assume a Bernoulli (binary) distribution for the response variable and calculate niche overlap (NO) between species i and j (on axis t) as:

$$NO_{ijt} = \min(p_{it}, p_{jt}) + \min(q_{it}, q_{jt}) \qquad \text{eqn 1}$$

where species i has probability p_i for 'success' (value 1) and $q_i = 1 - p_i$ for 'failure' (value 0). The proportion of individuals of species i with value 1 is the estimate of p_i. Similarly, p_j is estimated by the proportion of individuals of species j with value 1.

Categorical data

An example of categorical data is prey types in stomach content analysis (e.g. pelagic teleosts, demersal teleosts, invertebrates and chondrichthyans: Lucifora *et al.* 2008). There are K categories assumed to be equally available to species i. In an extension from binary data (two categories) to K categories, the niche overlap (NO) between species i and j (on axis t) is:

$$NO_{ijt} = \sum_{k=1}^{K} \min(p_{ikt} p_{jkt}) \qquad \text{eqn 2}$$

where p_{ik} is the proportional usage of category k by species i, assuming the sum of p_{ik} across all k categories equals 1. Similarly, species j has proportions p_{jk}.

For a graphical representation, see Appendix 1a.

Continuous data

Examples of continuous measures include many quantitative functional traits (e.g. specific leaf area: Beaumont & Burns 2009) and environmental covariates (e.g. salinity: Clarke & Allaway 1993). Using

kernel density estimations, Mouillot *et al.* (2005) converted finite data sets into continuous probability densities of flexible shape, avoiding two problems: (1) the loss of information involved in replacing continuous measurements with discrete categories; and (2) the unwarranted assumption of normality (or some other particular shape of distribution) if a single continuous distribution is fitted to the data. Density estimation by the kernel method (Silverman 1986) gives a smooth, flexible, nonparametric curve for a probability density function over the data points; however, the choice of the kernel bandwidth is a critical issue. Smaller kernel bandwidths concentrate the kernel function around the observed value, reducing estimated overlap, while larger bandwidths increase the width of the kernel, increasing estimated overlap (see discussions in: Stine & Heyse 2001; Mouillot *et al.* 2005). Silverman (1986) proposes an optional bandwidth based on the standard deviation of the data set and the population size when applying normal kernel density to normal data; however, as highlighted by Mouillot *et al.* (2005), further evaluation of the application of this bandwidth to nonparametric data is required.

For continuous data, niche overlap based on nonparametric kernel density functions (NO_K) on axis t is calculated as:

$$NO_{K_{i,j,t}} = 1 - \frac{1}{2} \int |f_{it}(x) - f_{jt}(x)| \mathrm{d}x \qquad \text{eqn 3}$$

where f_{it} and f_{jt} are the kernel population density functions for species i and j, respectively.

Modelling observed data with a mixture of normal distributions, which do not observe a restriction to positive values, may result in the extension of density curves to $x < 0$ when the data were positive measurements. Although this may be appropriate for data that can go below zero (e.g. minimum temperature), it is not appropriate for data that must be positive [e.g. fish standard length (SL)]. In the later case, Silverman (1986: section 2·10) suggests it is preferable to estimate the density of $\log(x)$. For a graphical representation, see Appendix 1b.

Ratio, proportion and percentage data

With appropriate transformations, ratio, proportion and percent data can be modelled in the same way as continuous data. Ratio data is a continuous positive measurement; therefore (as with continuous data), density estimation is appropriate. Ratio data requires one of two transformations depending on how the data is bounded. Ratio data that is bounded below by zero, but has no upper bounds (e.g. ratio of leaf area to leaf mass: see Appendix 1b) requires density estimation on $\log(x)$. Ratio data that is bounded below by zero and above by one (e.g. proportion data: for example, the proportion of tail length to total body length in lizards) requires density estimation on $\mathrm{logit}(x)$. This transformation prevents the density estimation overflowing the (0,1) bounds. The same approach is used for percentage data, which is bounded below by 0 and above by 100.

Count data

An example of count data is the number of prey items eaten by an individual during a period of observation (e.g. Jansen *et al.* 2002). For species i, there will be records from several individuals. Density estimations are created by taking a finite mixture of Poisson distributions, which is analogous to the mixture of normal distributions used in kernel density estimation for continuous data. Using nonparametric maximum likelihood estimation (NPMLE: see Norris & Pollock 1998), a mixture of finitely many Poisson distributions can be fitted to count data, with Akaike's Information Criterion (Akaike 1973) providing an objective choice of the number of components needed to

provide a good fit (analogous to bandwidth selection in kernel density estimation of continuous data).

For count data, the niche overlap (NO) for axis t is calculated as:

$$\mathrm{NO}_{i,j,t} = \sum_{x=0}^{x_{max}} \min\{p_{ixt}, p_{jxt}\} \qquad \text{eqn 4}$$

where the fitted distributions for species i and j have probabilities p_{ix}, p_{jx} respectively for values $x = 0, 1, 2, 3,...$, and x_{max} is set high enough for the minimum probability to be almost zero.

Electivity data

An example of electivity data is patterns of habitat use relative to availability in reef fish (e.g. McDermott & Shima 2006). Electivity scores are similar to proportion data (bounded below by 0 and above by 1), but they are not strictly continuous because there may be clusters of observations at 0 or at 1. If axis t is the usage of a resource, and all individuals of species i associate with that resource, the electivity score (e.g. Manly's Alpha: Manly, Miller & Cook 1972; Chesson 1978) is 1. If no individuals of species i associate with that resource, the electivity score is 0; however, between 0 and 1, with partial usage of the resource, the electivity measure is continuous. This implies there is a composite, or mixed, statistical distribution, neither fully continuous nor fully discrete. Such a distribution is not easily represented as a density. If a continuous curve is used for the probability density function on $0 < x < 1$, there should be infinite spikes at 0 and 1; similarly, the bar charts used in discrete distributions cannot accommodate continuous density functions. We note that representation as the cumulative distribution function is possible, $y = F(x) = Prob(X \le x)$. With discrete distributions, this rises from 0 to 1 in a step function, while a continuous distribution gives a continuously rising curve. A distribution with a mixture of discrete and continuous probability simply includes both steps and a continuous curve; however, it is not possible to display niche overlap on the cumulative probability graph. Instead, we illustrate such distributions in a triptych graph that has left and right panels to display the probabilities of 0 and 1 respectively, and a central panel for the continuous probability (see Appendix 1c). The density estimation for the central panel uses logit transformed data to avoid an overflow outside the lower bound of zero and the upper bound of one.

For species i, the estimated probability of 0 is the proportion of zeros, and the probability of 1 is the proportion of ones. The remaining observations have density estimation as for proportion data; however, the probabilities from the density estimation are downscaled to ensure the total probability (including values zero and one) equals one. The same procedure is used for species j. The combined overlap between species i and j in the three density estimates is then taken as the degree of niche overlap.

CONSTRUCTING UNIFIED INDICES OF NICHE OVERLAP AND THEIR GRAPHICAL REPRESENTATION

Having calculated overlap between species i and j for each data type in a mixed data set, unified measures of niche overlap may be obtained by averaging niche overlap (NO) between species i and j over each different axis t, where T is the number of axes:

$$\mathrm{NO}_{i,j} = \frac{1}{T}\sum_{t=1}^{T} \mathrm{NO}_{i,j,t} \qquad \text{eqn 5}$$

$\mathrm{NO}_{i,j}$ is 0 when the two distributions are completely disjoint, and is 1 when they exactly coincide.

Mouillot *et al.* (2005) also provide approaches for calculating niche overlap for traits that are related to similar functions. In this scenario, the degree of independence of each trait is discounted by downweighting highly correlated traits (see formulas 8–10: Mouillot *et al.* 2005).

The niche overlap measure may be viewed as a measure of association between species pairs, therefore, $d_{i,j} = 1 - \mathrm{NO}_{i,j}$ may be viewed as a measure of distance between species i and j. Using these distance measures, an $n \times n$ distance matrix $\mathrm{D} = (d_{i,j})$ can be constructed, where n is the number of species in the study. From this matrix, non-metric multidimensional scaling (nMDS), which reduces the dimensionality of the data below n (if justified by the stress levels in the scaling) can be used to graphically display, in two dimensions, niche relationships amongst a group of n species. Multidimensional scaling is only one of several multivariate techniques that could be used for dimension reduction (others include principal coordinates analysis – PCoA; principal component analysis – PCA; or linear discriminant analysis – LDA). If the assumption of a linear relationship between initial distances and distances in ordination space is justified, PCoA may be used. PCoA is advantageous in that it maximizes the linear correlation between initial distances in the dissimilarity matrix and distances in the ordination space.

DETECTING STATISTICAL DIFFERENCES IN NICHE OVERLAP

There are two alternative approaches to assessing differential use of niche space by multiple species. In the first approach, the relative position of two species within niche space is assessed to determine if their niches differ. The second approach examines the distribution of multiple species within niche space to determine if they are evenly distributed across, or clustered within, niche space (see Appendix 2).

Do the niches of two species differ?

A major objective of studies examining niche overlap is to determine if two (or more) species occupy the same niche. Even if two or more niches are identical, there will be some differences in the data purely by chance. To rule out detecting as different two niches that only differ due to sampling variation, niche comparisons between two species must be done statistically to determine whether the same probability distribution describes the niche of two (or more) species, or whether there is evidence of some difference. On any axis t, species i has n_i readings $\{x_{i1}, x_{i2}, ..., x_{in_i}\}$, which are realizations of the random variable X_{it}; similarly, species j has n_j readings $\{x_{j1}, x_{j2}, ..., x_{jn_j}\}$, from the random variable X_{jt}. The observed readings are used to estimate the parameters and hence the probability structure of each distribution; however, even if the random variables X_{it} and X_{jt} are the same, sampling variation will almost certainly cause the sampled data values to differ, giving $\mathrm{NO}_{ijt} < 1$. The question of interest is whether NO_{ijt} is sufficiently < 1 to provide evidence of niche differentiation. The same argument applies to the combined NO measure, averaged over all the axes.

Because the assumptions of equal variance and normality are unlikely to be met, we recommend analysis using null models and their associated permutation tests (Gotelli & Graves 1996; Gotelli 2000). Null models use randomization or permutation tests which do not rely on distributional assumptions (Manly 2007). Calculating pseudo-values of the test statistic that would arise if H_0 (no niche differentiation) were true generates the null distribution. This is achieved by permuting species labels over all species (i.e., average niche overlap over all species). The justification is that if all the species in

question occupy the same niche, the actual labelling of each species is irrelevant. This has the advantage of (1) retaining the total amount of each resource that is used; and (2) incorporating individual variation in resource use. If the set of species in the test is A, the test statistic is:

$$\overline{NO} = mean_{ij \in A, i < j} NO_{ij} \qquad \text{eqn 6}$$

with the mean niche overlap for all pairs of species (*i* and *j*) belonging in the set A (*i, j* 'in' A), with *i* the first species and *j* the second (*i < j*). If this is significantly higher than would be obtained under random assortment, there is niche overlap among the species within set A. The *P*-value is the proportion of pseudo-values less than the data-based value. If an overall niche difference exists between species *i* and *j*, details of which axes contribute to the difference may be obtained by comparing the data-based NO_{ijt} with the pseudo-values of NO_{ijt} and finding the *P*-value for each axis in turn; however, doing *t*-tests brings the problem of multiple comparisons, so an adjustment to the *P*-values is necessary to protect against false positives ('detecting' a difference which is not really there). We suggest a sequential Bonferroni adjustment (see Quinn & Keough 2005).

Are species within a community evenly distributed across niche space?

In community ecology, the distribution of species across niche space is often assessed to determine if species coexist via resource partitioning. Given *i* species and *t* axes, for each axis, there are $\frac{i(i-1)}{2}$ niche overlap values (say $x_1, x_2, ..., x_n$) with an associated mean ($\bar{x}$) and variance (s^2). To evaluate how species are distributed across niche space we can consider two extreme cases of the variance in relation to the mean. A minimum variance of zero occurs if the individual species' niches are evenly spaced across niche space, with equal mean niche overlap for each adjacent pair of species. The actual value of mean overlap will be set by the boundaries of niche space. This represents the maximum possible separation of species on this niche axis. Alternatively, maximum variance occurs when all niche overlaps are either 0 or 1, a proportion being at 1 ($p = \bar{x}$) and the remaining proportion being at zero ($1 - p$). The maximum variance is $\bar{x}(1 - \bar{x})$ or $p(1 - p)$. This case represents species clustering (the opposite of species separation). Between these two extremes, there will be a variance associated with random distribution of the species on the niche axis, with neither even spacing nor clustering occurring (see Appendix 3). Therefore, the proportion of maximum variance (i.e., the coefficient of heterogeneity: see Pledger 2005) can be used as a measure of even spacing vs. clustering:

$$\eta = \frac{s^2}{\bar{x}(1 - \bar{x})} = \frac{s^2}{p(1 - p)} \qquad \text{eqn 7}$$

A value near 0 indicates even spacing, while a value near 1 indicates clustering. The shape and position of the curve of random association is difficult to calculate; however, by calculating η for pseudo-values generated by permuting species labels so that all species are randomly placed in this niche dimension, a test statistic can be constructed to determine if the pattern in the data is random (the null model), clustered (η near 1) or evenly spaced (η near 0). Having calculated η for each niche axis, examining the minimum and maximum values across all *t* niche axis will indicate if clustering occurs in one axis while there is even spacing in another. Finally, the average of η over the *t* niche axis gives an overall measure of clustering vs. even spacing.

CORAL REEF FISH EXAMPLE

To illustrate the use of overlap indices on multiple niche axes generated from data of different types, we use data from individuals < 25 mm SL from five fish species found between the shore and barrier reefs crest of the northern lagoon of Moorea, French Polynesia (17°30′S, 149°50′W): *Gomphosus varius* (Lacepéde, 1801); *Pseudocheilinus hexataenia* (Bleeker, 1857), *Scarus sordidus* (Forsskål, 1775), *Thalassoma hardwicke* (Bennett, 1830) and *T. quinquevittatum* (Lay & Bennett, 1839). For each species, we quantified local realized niche space along three axes, incorporating three different types of data: (1) Pectoral fin aspect ratio (ratio data bounded below by 0); (2) habitat association (electivity data); and (3) distributions across the lagoon (categorical data). Pectoral fin aspect ratio was quantified as leading edge^2 per area (following Wainwright, Bellwood & Westneat 2002). Pectoral fin aspect ratio is related to manoeuvrability, swimming speed and efficiency of locomotion (Walker & Westneat 1997; Wainwright, Bellwood & Westneat 2002). The use of individuals of different sizes can introduce allometric bias into shape analysis (McCoy *et al.* 2006). We examined each species for allometric bias by regressing pectoral fin aspect ratio against SL, resulting in linear regressions described by $Y = a + \beta_X$. We tested the significance of the slope (β) obtained in each regression in order to determine if the *Y* variables showed isometric (β = 0) or allometric (β ≠ 0) growth in relation to SL. We found no evidence against isometry (β ≠ 0, *P* > 0·05 in all cases). Pectoral fin aspect ratio was determined for 19, 14, 17, 16 and 15 individuals for *G. varius*, *P. hexataenia*, *S. sordidus*, *T. hardwicke* and *T. quinquevittatum*, respectively.

Habitat association was quantified as Manly's Alpha (Manly, Miller & Cook 1972; Chesson 1978), the proportion of fish at a given reef, on a given habitat type, relative to the probability of the fish associating with that habitat type under conditions of random assortment. In March 2005, we established three transects within Moorea's northern lagoon. Each transect was *c*. 800 m in length, and ran from shore to the barrier reef crest. We selected five equally spaced sites along each transect. Within each site, we haphazardly selected 16 patch reefs (hereafter reefs) of similar size [*n* = 240 (= 3 × 5 × 16) reefs: average surface area = 7·98 m^2, SD = 3·79; average height = 0·83 m, SD = 0·22]. For each reef, and a 1-m halo surrounding the reef, we visually estimated percent cover of six substrate categories: *Porites* massive; branching corals; macroalgae; bare habitat; other and rubble. We also recorded the presence of all individuals of the study species < 25 mm SL, and noted the substrate each individual was associated with at the time it was first observed. We made all observations between 08:00 h and 16:00 h (time of peak activity of benthic fishes) from 27 May to 6 June 2005. Mean overlap in habitat association was calculated separately for each of six habitat categories on 240 reefs, and then combined into a single index of habitat association using formula 5 above.

Using the data set from (2) above, we calculated distribution across the lagoon as the density of each species on each reef at the five sites between the barrier reef crest and shore (crest, offshore, centre, inshore, shore).

For each axis (pectoral fin aspect ratio, habitat association and distribution within the lagoon), we calculated niche overlap between the five species using the appropriate modelling technique. We then constructed unified measures of mean niche overlap and associated variance over the three axes using equation 5 above. We used Kruskal's nMDS to graphically display results (Cox & Cox 2001). We constructed test statistics from 1000 permutations of species labels using null models to determine if local realized niche space differs

between pairs of species, and if the local realized niches of all species are evenly distributed across niche space.

ALPINE PLANT EXAMPLE

We also illustrate the use of overlap indices using data from five plant species found along an elevational gradient in the Nelson Lakes National Park of New Zealand's South Island (41°49′S, 172°50′E): *Coprosma foetidissima* (Forst & Forst, 1776); *Coprosma microcarpa* (Hooker, 1853); *Coprosma pseudociliata* (Jane, 2005); *Coprosma pseudocuneata* (Garnock-Jones & Elder 1996) and *Coprosma rhamnoides* (Cunningham 1839). For each species, we quantified local realized niche space along two axes incorporating two different types of data: (1) Distribution along an elevational gradient (continuous data: 26 elevations; 620, 655, 690, 730, 770, 795, 820, 860, 885, 925, 960, 1000, 1030, 1055, 1075, 1100, 1125, 1150, 1195, 1230, 1270, 1310, 1345, 1385, 1420 and 1460 m a.s.l) taken as the number of each species in five 10 m^2 randomly situated plots at each elevation. (2) Specific leaf area (ratio data bounded below by 0), defined as the ratio between leaf blade area and its dry weight (Garnier *et al.* 2001). Although specific leaf area may not be an appropriate measure of niche space, we use it here as an exemplar. Specific leaf area is representative of different life-history strategies, ranging from low-cost, short-lived leaves that provide a quick return on their energetic investment, to long-lived, energetically costly leaves that offer slower returns relative to energetic investment (see Whitfield 2006). Specific leaf area was determined for 4, 8, 5, 13 and 5 individuals for *C. foetidissima*, *C. microcarpa*, *C. pseudociliata*, *C. pseudocuneata* and *C. rhamnoides*, respectively. For each axis (distribution along an elevational gradient and specific leaf area), we calculated local realized niche overlap between the five species using the appropriate modelling technique, and then constructed multivariate measures of mean niche overlap using equation 5 above. We used nMDS to graphically display results. We constructed test statistics from 1000 permutations of species labels using null models to determine if pairs of species occupied different niche space, and if all species were evenly distributed across niche space.

All statistical analysis was conducted in R 2·11·1 (R Development Core Team 2010). We conducted nMDS using the algorithm in the R package MASS, command isoMDS (R Development Core Team 2010). The stress function used in nMDS was the square root of the ratio of the sum of squared differences between the input distances and those of the configuration to the sum of configuration distances squared. Appendix 4 contains R-code and data sets for conducting the reef fish and alpine plant analyses.

Results

CORAL REEF FISH EXAMPLE

Of the five fish species, the local realized niches of *P. hexataenia* and *S. sordidus* (overlap = 0·470; Table 1) were most dissimi-

lar, and local realized niches of *P. hexataenia* and *T. hardwicke* were most similar (overlap = 0·811; Table 1). *Scarus sordidus* did not have the same local realized niche as any of the other species (Table 1, Fig. 1). *Gomphosus varius*, *P. hexataenia* and *T. hardwicke* had similar local realized niches (Table 1, Fig. 1). Although the local realized niche of *T. quinquevittatum* was similar to *G. varius*, it differed from *P. hexataenia* and *T. hardwicke* (Table 1, Fig. 1). The local realized niches of the five species displayed significant clustering in their distribution across niche space (Table 2, Fig. 1).

ALPINE PLANT EXAMPLE

Of the five alpine plants, the local realized niches of *C. pseudocuneata* and *C. foetidissima* were most dissimilar (overlap = 0·101, Table 3). Conversely, the local realized niches of *C. foetidissima* and *C. rhamnoides* were most similar (overlap = 0·582, Table 3), and these species did not differ in their local realized niches (Fig. 2). *Coprosma rhamnoides* and *C. foetidissima* had similar local realized niches (Table 3, Fig. 2). The five species displayed significant clustering in their distribution across niche space (Table 4, Fig. 2).

Discussion

A major objective of studies examining niche overlap is to determine if two or more species occupy the same niche. This requires first identifying the position in niche space occupied by each species, and then determining if two or more species occupy the same portion of niche space. To the best of our knowledge, there are currently no formal indices capable of describing niche space using multiple axes described by different data types. Although multivariate niche analysis is a well established field (e.g. Green 1971; Harner & Whitmore 1977; Mahdi, Law & Willis 1989; Litvak & Hansell 1990; Pappas & Stoermer 1997), this article presents a novel analytical approach, capable of providing a composite measure of niche overlap, derived from a wide range of data types: binary, categorical, various continuous data types, and electivity data. Also presented are novel permutation tests that assess if pairs of species occupy the same niche space, and if all species are evenly distributed across niche space.

Of the five coral reef fishes examined in the formal analysis of niche overlap, four were wrasses and one (*S. sordidus*) was a parrotfish. We were therefore not surprised to find that the local realized niche of *S. sordidus* differed from that of the wrasse species, which have similar local realized niches.

	Pseudocheilinus hexataenia	*Scarus sordidus*	*Thalassoma hardwicke*	*Thalassoma quinquevittatum*
Gomphosus varius	0·730 (0·155)	**0·585 (0·194)**	0·775 (0·137)	0·785 (0·062)
P. hexataenia		**0·470 (0·376)**	0·811 (0·011)	**0·569 (0·160)**
S. sordidus			**0·512 (0·345)**	**0·570 (0·161)**
T. hardwicke				**0·632 (0·216)**

Table 1. Mean (SD) niche overlap between five species of reef fishes incorporating three functional traits: (1) pectoral fin aspect ratio (ratio data bound by 0 and 1); (2) habitat association (electivity data) and (3) distance from shore (categorical data)

Species pairs occupying statistically different niches, as identified by null model tests, are indicated in bold ($P < 0·01$).
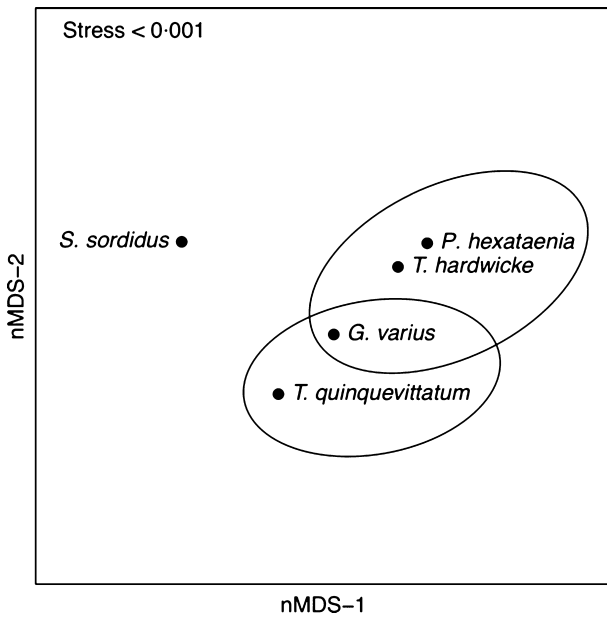
Fig. 1. Interspecific similarities in 'unified' niche overlap between five species of reef fishes incorporating three functional traits: (1) pectoral fin aspect ratio (measurement data); (2) habitat association (electivity data) and (3) distance from shore (categorical data). Similarities are represented graphically as non-metric multi-dimensional scaling. Ellipses encircle species occupying niches that were not identified as significantly different using null model tests.

Table 2. Measures of the evenness vs. clustering of five species of reef fishes across niche space incorporating three functional traits: (1) pectoral fin aspect ratio (ratio data bound by 0 and 1); (2) habitat association (electivity data); and (3) distance from shore (categorical data); with the average coefficient of heterogeneity over the three functional traits giving an overall measure of clustering vs. even spacing (see *Methods*)

	Pectoral fin aspect ratio	Habitat association	Distance from shore	Overall
Evenness test	> 0·999	0·985	0·980	0·992
Clustering test	< 0·001	0·015	0·020	**0·008**

Statistically significant evenly spaced or clustered distributions, as identified by null model tests, are indicated in bold ($P < 0.05$).

The local realized niches of *S. sordidus* and *P. hexataenia* were most dissimilar. These species have distinct functional niches; the former over-associates with *Porites rus*, *Turbinaria ornata* and coral rubble (Geange 2010), feeding primarily on algae
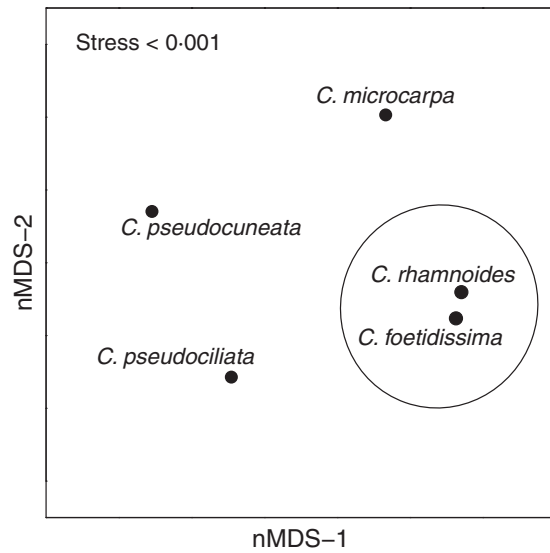

Fig. 2. Interspecific similarities in 'unified' niche overlap between five alpine plant species incorporating two functional traits: (1) surface leaf area (continuous data) and (2) elevation (categorical data). Similarities are represented graphically as non-metric multi-dimensional scaling. Ellipses encircle species occupying niches that were not identified as significantly different using null model tests.

(Chen 2002). *Pseudocheilinus hexataenia* over-associates with the structurally complex hard corals *Pocillopora verrucosa* and *Porites rus* (Geange 2010), feeding on copepods, amphipods, gastropods, shrimps and polychaete worms (Harmelin-Vivien 1979). Conversely, the local realized niches of *P. hexataenia* and *T. hardwicke* were most similar. Like, *P. hexataenia*, *T. hardwicke* feeds primarily on small invertebrates (Myers 1999) and over-associates with *Pocillopora verrucosa* and *Porites rus* (Geange 2010).

In the analysis of alpine plant species, lowest overlap occurred between *C. pseudocuneata* and *C. foetidissima*. These species have very different elevational distributions. *Coprosma pseudocuneata* occupies higher montane to subalpine forest while *C. foetidissima* is restricted to lowland to lower montane forest and shrubland (Allan 1961). Additionally, the leaves of *C. pseudocuneata* are thick and coriaceous, while those of *C. foetidissima* are membranous subcoriaceous leaves (Allan 1961), suggesting that these species have very different energetic investments in leaf production. Highest niche overlap occurred between *C. foetidissima* and *C. rhamnoides*, both of which were found at lower elevations (the former inhabits coastal to sub-alpine forests: Allan 1961). The marked

Table 3. Mean (SD) niche overlap between five alpine plant species incorporating two functional traits: (1) elevation (continuous data); and (2) surface leaf area (ratio data bounded below by 0)

	Coprosma microcarpa	*Coprosma pseudociliata*	*Coprosma pseudocuneata*	*Coprosma rhamnoides*
Coprosma foetidissima	**0·363 (0·213)**	0·358 (0·470)	**0·101 (0·034)**	0·582 (0·094)
C. microcarpa		0·139 (0·151)	0·303 (0·250)	**0·413 (0·139)**
C. pseudociliata			0·478 (0·481)	0·310 (0·438)
C. pseudocuneata				**0·121 (0·129)**

Species pairs occupying statistically different niches, as identified by null model tests, are indicated in bold ($P < 0.01$).

Table 4. Measures of the evenness vs. clustering of five alpine plant species across niche space incorporating two functional traits: (1) Distribution along an elevational gradient (continuous data); and (2) Specific leaf area (ratio data bounded below by 0); with the average coefficient of heterogeneity over the two functional traits giving an overall measure of clustering vs. even spacing (see Methods)

	Elevational gradient	Surface leaf area	Overall
Evenness test	0·872	> 0·999	> 0·999
Clustering test	**0·128**	**< 0·001**	**< 0·001**

Statistically significant evenly spaced or clustered distributions, as identified by null model tests, are indicated in bold ($P < 0.05$).

similarities in elevational distribution of these species appears to result in similar local realized niches.

Deciding what a resource is, how each resource is subdivided, and how many resources to include in analyses of niche overlap is often subjective. Each of these decisions can have large effects on how species relate to one another in modelled niche space. For example, if microhabitats were subdivided into two categories (alive vs. dead coral), two species occupying live coral would have an overlap of 1; however, if these two species use different live coral microhabitats they would have an overlap of 0 if the different microhabitats were distinguished. Further, when multiple axes are included, some axes may be considered more important than others; this raises the issue of weighting. For example, if two fish species eat the same food, but one is nocturnal while the other is diurnal, it may be appropriate to down-weight the axis describing food source relative to the axis describing period of activity (see formulas 8–10: Mouillot *et al.* 2005). Unfortunately, there are no general rules for determining which axis to include and how it should be subdivided or weighted; these issues need to be informed by the researcher's knowledge of the specific study system and their relevance to the coexistence of the species in question.

We used null models to compare niche overlap between species. Although null models are viewed as a useful tool for revealing pattern (or lack thereof) in natural communities, the choice of appropriate null models has been highlighted as a critical issue in null model analysis (Gotelli 2001). In practice, it is often difficult to decide what constitutes a truly null, purely statistical model. For example, it is difficult to be certain that some biological interactions that may have shaped the data are not woven into the null model (Harvey *et al.* 1983). Lawlor (1980) developed four randomization algorithms that are commonly used to construct null models. These algorithms differ in whether utilizations are reshuffled or replaced by a random number, and in whether the zeros in the matrix are retained or not. Both decisions have implications for the structure of the null community, and affect the power of the test (Gotelli & Graves 1996). The properties of these randomization techniques have previously been explored (see Winemiller & Pianka 1990; Kobayashi 1991), and have identified Lawlor's (1980) RA3 as the best existing algorithm for use in resource overlap null models. Like Lawlor's RA3 algorithm, our approach scrambles zeros, retaining the original total amount of each resource utilized; however, there is a fundamental difference between our randomization procedure and RA3. Lawlor's RA3 reallocates pooled resource usages for each species, whereas our approach reallocates resource use for each individual. Because our approach does not assume all individuals of a species act alike (i.e., incorporates individual variation), it seems most powerful for tests determining how individuals of different species differ in their distribution along a niche axis.

Although null models can be used to establish whether observed niche overlap is more or less than expected by chance, it is still difficult to infer the mechanism(s) responsible for such patterns. For example, high niche overlap may reflect intense competition for shared resources or, alternatively, a surplus of resources and the absence of competition (Glasser & Price 1988). Both scenarios have been revealed in experimental field studies of competition (Schoener 1983). It is not our intention to provide insights into the underlying causes of a given degree of niche overlap, but to outline and illustrate our approach in calculating unified niche overlap incorporating measures of different types.

Although we illustrated the implementation of our unified indices of niche overlap by examining interspecific overlap, it is also worth noting that our methodology can also be used to test for intraspecific temporal or spatial niche shifts by quantifying differences in niche overlap for a single species between different places or times. With this new approach, we have developed the ideas of Slobodchikoff & Schulz (1980) to calculate overlap across several niche dimensions incorporating binary, categorical, ratio, proportion, percent, count and electivity data. We have also devised null model tests to assess statistical differences in niche overlap between pairs of species, and if species within a community are evenly distributed across, or clustered within niche space.

Acknowledgements

References

Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory* (eds B.N. Petrov & F. Csaki), pp. 267–281. Academiai Kiado, Budapest.

Allan, H.H. (1961) *Flora of New Zealand*. Government Printer, Wellington.

Beaumont, S. & Burns, K.C. (2009) Vertical gradients in leaf trait diversity in a New Zealand forest. *Trees*, **23**, 339–346.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.

Chen, L.-S. (2002) Post-settlement diet shift of *Chlorurus sordidus* and *Scarus schlegeli* (Pisces: Scaridae). *Zoological Studies*, **41**, 47–58.

Chesson, J. (1978) Measuring preference in selective predation. *Ecology*, **59**, 211–215.

Clarke, P.J. & Allaway, W.G. (1993) The regeneration niche of the grey mangrove (*Avicennia marina*): effects of salinity, light and sediment factors on establishment, growth and survival in the field. *Oecologia*, **93**, 548–556.

Cody, M. (1975) Towards a theory of continental species diversities: bird distributions over Mediterranean habitat gradients. *Ecology and Evolution of Communities* (eds M. Cody & J. Diamond), pp. 214–257. Harvard University Press, Cambridge.

Cox, T.F. & Cox, M.A.A. (2001) *Multidimensional Scaling*. 2nd edn. Chapman & Hall/CRC, Florida, USA.

Day, T. & Young, K. (2004) Competitive and facilitative evolutionary diversification. *BioScience*, **54**, 101–109.

Dolédec, S., Chessel, D. & Gimaret-Carpentier, C. (2000) Niche separation in community analysis: a new method. *Ecology*, **81**, 2914–2927.

Dueser, R.D. & Shuggart, H.H. (1979) Niche pattern in a forest-floor small-mammal fauna. *Ecology*, **60**, 108–118.

Elton, C. (1927) *Animal Ecology*. Sidgwick and Jackson, London.

Garnier, E., Shipley, B., Roumet, C. & Laurent, G. (2001) A standardized protocol for the determination of specific leaf area and leaf dry matter content. *Functional Ecology*, **15**, 688–695.

Geange, S.W. (2010) Effects of larger heterospecifics and structural refuge on the survival of a coral reef fish, *Thalassoma hardwicke*. *Marine Ecology Progress Series*, **407**, 197–207.

Glasser, J.W. & Price, H.J. (1988) Evaluating expectations deduced from explicit hypotheses about mechanisms of competition. *Oikos*, **51**, 57–70.

Gotelli, N.J. (2000) Null model analysis of species co-occurrence patterns. *Ecology*, **81**, 2606–2621.

Gotelli, N.J. (2001) Research frontiers in null model analysis. *Global Ecology & Biogeography*, **10**, 337–343.

Gotelli, N.J. & Graves, G.R. (1996) *Null Models in Ecology*. Smithsonian Institute Press, Washington D.C.

Green, R.H. (1971) A multivariate statistical approach to the Hutchinsonian niche: bivalve molluscs of Central Canada. *Ecology*, **52**, 544–556.

Gregory, S.D. & Macdonald, D.W. (2009) Prickly coexistence or blunt competition? Opuntia refugia in an invaded rodent community. *Oecologia*, **159**, 225–236.

Grinnell, J. (1917) The niche relationships of the California Thrasher. *The Auk*, **34**, 427–433.

Harmelin-Vivien, M.L. (1979) Ichtyofaune des récifs coralliens de Tuléar (Madagascar). Ecologie et relations trophiques. In. Université Aix-Marseille II, France, p. 165.

Harner, E.J. & Whitmore, R.C. (1977) Multivariate measures of niche overlap using discriminant analysis. *Theoretical Population Biology*, **12**, 21–36.

Harvey, P.H., Colwell, R.K., Silvertown, J.W. & May, R.M. (1983) Null models in ecology. *Annual Review of Ecology and Systematics*, **14**, 189–211.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

Jansen, P.A., Slettvold, H., Finstad, A.G. & Langeland, A. (2002) Niche segregation between Arctic char (*Salvelinus alpinus*) and brown trout (*Salmo trutta*): an experimental study of mechanisms. *Canadian Journal of Fisheries and Aquatic Science*, **59**, 6–11.

Kobayashi, S. (1991) Interspecific relations in forest floor Coleopteron assemblages: niche overlap and guild structure. *Researches on Population Ecology*, **33**, 345–360.

Lawlor, L.R. (1980) Structure and stability in natural and randomly constructed competitive communities. *The American Naturalist*, **116**, 394–408.

Litvak, M.K. & Hansell, R.I.C. (1990) A community perspective on the multidimensional niche. *Journal of Animal Ecology*, **59**, 931–940.

Lucifora, L.O., García, V.B., Menni, R.C., Escalante, A.H. & Hozbor, N.M. (2008) Effects of body size, age and maturity stage on diet in a large shark: ecological and applied implications. *Ecological Research*, **24**, 109–118.

MacArthur, R. & Levins, R. (1967) Limiting similarity, convergence, and divergence of coexisting species. *American Naturalist*, **101**, 377–385.

Mahdi, A., Law, R. & Willis, A.J. (1989) Large niche overlaps among coexisting plant species in a limestone grassland community. *Journal of Ecology*, **77**, 386–400.

Manly, B.F.J. (1994) Ecological statistics. *Handbook of Statistics*, Vol. 12, Environmental Statistics (eds G.P. Patil & C.R. Rao), pp. 307–376. Elsevier Science, Amsterdam.

Manly, B.F.J. (2007) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 3rd edn. Chapmann and Hall, London.

Manly, B.F.J., Miller, P. & Cook, L.M. (1972) Analysis of a selective predation experiment. *The American Naturalist*, **106**, 719–736.

McCoy, M.W., Bolker, B.M., Osenberg, C.W., Miner, B.G. & Vonesh, J.R. (2006) Size correction: comparing morphological traits among populations and environments. *Oecologia*, **148**, 547–554.

McDermott, C.J. & Shima, J.S. (2006) Ontogenetic shifts in microhabitat preference of the temperate reef fish *Forsterygion lapillum*: implications for population limitation. *Marine Ecology Progress Series*, **320**, 259–266.

Moilanen, A. & Nieminen, M. (2002) Simple connectivity measures in spatial ecology. *Ecology*, **83**, 1131–1145.

Mookerji, N., Weng, Z. & Mazumder, A. (2004) Food partitioning between coexisting Atlantic Salmon and brook trout in the Saint-Marguerite river ecosystem, Quebec. *Journal of Fish Biology*, **64**, 680–694.

Mouillot, D., Stubbs, W., Faure, M., Dumay, O., Tomasini, J., Wilson, J. & Chi, T. (2005) Niche overlap estimates based on quantitative functional traits: a new family of non-parametric indices. *Oecologia*, **145**, 345–353.

Myers, R.F. (1999) *Micronesian Reef Fishes: A Field Guide for Divers and Aquarists*. Coral Graphics, Guam.

Norris, J.L. & Pollock, K.H. (1998) Non-parametric MLE for Poisson species abundance models allowing for heterogeneity between species. *Environmental and Ecological Statistics*, **5**, 391–402.

Olden, J.D., Poff, N.L. & Bestgen, K.R. (2006) Life-history strategies predict fish invasions and extirpations in the Colorado River Basin. *Ecological Monographs*, **76**, 25–40.

Pappas, J.L. & Stoermer, E.F. (1997) Multivariate measure of niche overlap using canonical correspondence analysis. *Ecoscience*, **4**, 240–245.

Pianka, E.R. (1973) The structure of lizard communities. *Annual Review of Ecology and Systematics*, **4**, 53–74.

Pledger, S. (2005) The performance of mixture models in heterogeneous closed population capture–recapture. *Biometrics*, **61**, 868–873.

Pledger, S. & Geange, S. (2009) Niche overlap: a unified definition and analysis for data of different types. *Research Report Series*. School of Mathematics, Statistics and Computer Science, Victoria University of Wellington, New Zealand.

Quinn, G.P. & Keough, M.J. (2005) *Experimental Design and Data Analysis for Biologists*, 4th edn. Cambridge University Press, Cambridge, UK.

R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schoener, T. (1983) Field experiments on interspecific competition. *American Naturalist*, **122**, 240–285.

Silverman, B.W. (1986) *Density Estimation for Statistics Ad Data Analysis*. Chapman and Hall, London.

Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*, **19**, 605–611.

Silvertown, J., Dodd, M., Gowing, D., Lawson, C. & McConway, K. (2006) Phylogeny and the hierarchical organization of plant diversity. *Ecology*, **87**, S39–S49.

Slobodchikoff, C.N. & Schulz, W.C. (1980) Measures of niche overlap. *Ecology*, **61**, 1051–1055.

Smith, F.A., Brown, J.H., Haskell, J.P., Lyons, S.K., Alroy, J., Charnov, E.L., Dayan, T., Enquist, B.J., Ernest, S.K.M., Hadly, E.A., Jones, K.E., Kaufman, D.M., Marquet, P.A., Maurer, B.A., Niklas, K.J., Porter, W.P., Tiffney, B. & Willig, M.R. (2004) Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *American Naturalist*, **163**, 672–691.

Stine, R. & Heyse, J. (2001) Non-parametric estimates of overlap. *Statistics in Medicine*, **20**, 215–236.

Sugihara, G., Bersier, L.F., Southwood, T.R.E., Pimm, S.L. & May, R.M. (2003) Predicted correspondence between species abundances and dendrograms of niche similarities. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 5246–5251.

Wainwright, P.C., Bellwood, D.R. & Westneat, M.W. (2002) Ecomorphology of locomotion in labrid fishes. *Environmental Biology of Fishes*, **65**, 47–62.

Walker, J.A. & Westneat, M.W. (1997) Labriform propulsion in fishes: kinematics of flapping aquatic flight in the bird wrasse *Gomphosus varius* (Labridae). *Journal of Experimental Biology*, **200**, 1549–1569.

Whitfield, J. (2006) The cost of leafing. *Nature*, **444**, 539–541.

Winemiller, K.O. & Pianka, E.R. (1990) Organization in natural assemblages of desert lizards and tropical fishes. *Ecological Monographs*, **60**, 27–55.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix 1. Overlapping probability distributions constructed using three different data types.

Appendix 2. Species distributions across niche space.

Appendix 3. Variance of niche overlaps in relation to the mean.

Appendix 4. Source code and three examples for the calculation of multivariate niche overlap incorporating data of different types, with associated null model tests of (i) differential use of niche space by multiple species; and (ii) even distribution of species across niche space.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.