2-2010

Selection Ratios on Community Aggregated Traits to Estimate Ecological Filters Imposed on Species by Sites

Nathanael I. Lichti

Michael T. Murphy
Portland State University, murphym@pdx.edu

Citation Details

Lichti, Nathanael I., and Michael T. Murphy. Selection ratios on community aggregated traits estimate ecological filters imposed on species by sites. Ecology 91.2 (2010): 347-354.

Ecology, 91(2), 2010, pp. 347–354

Selection ratios on community aggregated traits estimate ecological filters imposed on species by sites

Nathanael I. Lichti[1] and Michael T. Murphy

Department of Biology, Portland State University, P.O. Box 751, Portland, Oregon 97207-0751 USA

Abstract. Variation in community structure is mediated by interactions between species traits and a site's environmental characteristics. Previously, data on community composition at sites has been employed to correlate trait and environmental variables (e.g., RLQ analysis) and to predict community-level expression of quantitative traits (i.e., community aggregated traits). Here, we demonstrate that the selection ratio, a method originating in animal resource selection studies, can estimate the ecological filters that site conditions impose on species traits by combining observed community aggregated traits with null models of species availability. This flexible, nonparametric approach expresses the filter at each site as a probability density function for the selection of individuals possessing a given trait value. By doing so, it generalizes the community aggregated trait concept to include categorical as well as continuous traits and allows for both intraspecific variation in trait expression and differences in species availability among sites. The resulting site-level filter functions can be related to environmental covariates by standard statistical approaches (e.g., regression). The method complements existing techniques for analyzing trait–environment interactions in community ecology.

Key words: assembly theory; community aggregated trait; community ecology; ecological filter; fourth-corner problem; selection ratio; trait–environment interaction.

INTRODUCTION

Natural selection (Darwin 1859) and niche theory (Hutchinson 1957) both depend on the interaction between environmental conditions and organisms' morphologic, physiologic, behavioral, and life-history characteristics. Until recently, however, ecologists have possessed few tools to quantify or test for relationships between specific environmental variables and species traits (Legendre and Legendre 1998). Legendre et al. (1997) labeled this apparent paradox as the fourth-corner problem: although a trait of interest (for example, drought tolerance in plants) may be shared to varying degrees by many species in a community, connections between environmental variables and traits may be obscured by other factors that affect species abundances at sites, such as competition.

To solve the fourth-corner problem, data on species abundances or occurrences and their traits must be combined into a site-level index of trait expression (Legendre et al. 1997). In quantitative community ecology, RLQ analysis links traits and sites in a three-way ordination by performing a double inertia analysis of two arrays, R (a site–environment table) and Q (a species–trait table), with a link expressed by a contingency table, L, of species abundances at sites (Dolédec et al. 1996). The analysis yields correlation coefficients for traits and environmental variables (Dray and Legendre 2008). A simpler approach has arisen in ecosystem ecology, where interest often lies in predicting the aggregated value of specific traits for a community, rather than estimating trait-environment correlations per se. The community aggregated trait (CAT) for a site is the mean value of a continuous trait expressed by each species at the site, weighted by the species' relative abundances (Garnier et al. 2004). CAT scores are straightforward to calculate and to model, but do not currently accommodate categorical traits or intraspecific variation in expression (Garnier et al. 2004, Shipley et al. 2006).

Community assembly theory posits that composition at a given time and place results from the application of a series of ecological filters to a larger, regional species pool (Weiher and Keddy 1995, Poff 1997, Díaz et al. 2007). These filters can be conceived of as functions for the probability that a given species will be found at a site, conditioned on the site's and species' respective characteristics. Community aggregated traits give a point estimate of the filter at a site, and a trait–environment correlation describes the filter's mean change along an environmental gradient. An estimate of the entire filter function for a trait therefore provides the information needed to evaluate both trait–environment interactions and CAT values. However, correct

---

Manuscript received 29 September 2008; revised 16 July 2009; accepted 5 August 2009. Corresponding Editor: B. J. Cardinale.

[1] Current address: Department of Forestry and Natural Resources, Purdue University, 715 W. State Street, West Lafayette, Indiana 47907 USA. E-mail: nlichti@purdue.edu

filter estimation requires that we know which species were available to colonize a site as well as the species found at the site. As typically implemented, both RLQ analysis and CAT assume that local communities all arise from the same regional pool. This assumption may not always be appropriate. For instance, the distribution of source populations may affect propagule availability.

We propose that differences among species pools can be accounted for by converting raw CAT scores into selection ratios. To do so, we generalize the CAT concept to include the distribution of trait expression at a site, instead of simply the mean. This allows for both categorical trait variables and intraspecific variation in trait expression. We then combine the CAT distribution with a null model that describes species pool effects (e.g., Kelt et al. 1995). The resulting distributions quantify the selective filters acting on a trait as site-specific probability functions which can be related to environmental covariates by linear regression, generalized linear modeling, quantile regression, or other standard methods depending on the researcher's analytical goals and distributional assumptions.

BACKGROUND

Selection ratios have been used for decades to study resource selection by animals and are covered authoritatively by Manly et al. (2002). To our knowledge, they have never been used to quantify trait–environment interactions. Additional discussions of the resource selection literature, including more advanced methods and areas of current development, are available in Manly et al. (2002) and in a special section of *The Journal of Wildlife Management* (Strickland and McDonald 2006). They may also be of interest in the context of community ecology.

Savage (1931) introduced selection ratios as the proportion of a resource used by an individual or a population divided by the proportion available in the environment. For example, if captive squirrels in an experiment on habitat preference are presented with equal quantities of four cover types, the estimated selection ratio for squirrel i and cover type j ($\hat{w}_{ij}$) compares the proportion of time the squirrel spends in habitat j (p_{ij}) to the proportion of the enclosure made up of habitat j and thus available for use (π_{+j}, which equals 0.25 in this case):

$$\hat{w}_{ij} = p_{ij}/\pi_{+j} = p_{ij}/0.25. \tag{1}$$

If the squirrel uses cover j in proportion to its availability, the ratio equals 1. Values significantly >1 indicate a preference, and values <1 indicate avoidance. To determine whether any particular $\hat{w}_{ij}$ deviates significantly from 1, we require a predictive distribution for $\hat{w}_{ij}$ under the assumption that $\mathrm{E}(\hat{w}_{+j}) = 1$. This distribution can be estimated by bootstrap simulation from a null model of availability (see Appendix).

Since the upper bound on $\hat{w}_{ij}$ is undetermined (it depends on π_{+j} and therefore on the specific data set in question), Manly (1993) suggests using the standardized selection index ($\hat{B}_{ij}$) instead:

$$\hat{B}_{ij} = \frac{\hat{w}_{ij}}{\sum_{j=1}^{n} \hat{w}_{ij}}. \tag{2}$$

The standardized indices can be interpreted as the probability that any item selected by i will belong to category j. The set $\hat{B}_i : (\hat{B}_{i1}, \ldots, \hat{B}_{in})$ is a probability mass function for resource selection conditioned on individual identity.

Manly et al. (2002) present numerous specialized variations on these themes, including formulations for estimated available proportions (i.e., $\hat{\pi}_j$), separate availability estimates for individuals ($\hat{\pi}_{ij}$), and population-level selection ratios. These variations provide ways to fine-tune calculations to specific sampling scenarios and objectives, and may be applicable in many studies of trait–environment interaction.

TRAIT SELECTION BY SITES

To motivate the application of selection ratios to trait–environment interactions, we reverse the idea that organisms choose their environments and instead consider selection from an evolutionary perspective. Following the species-pool–ecological-filter model, we assume that the species occurring at any point in space are drawn from a regional pool with expected relative frequencies (π_{ij}) equal to their proportional abundance in that pool (Kelt et al. 1995). The species pool can be the same for all points (hereafter *sites*), but this is not required. Nonrandom variation among local communities within a pool arises when the environment at different sites filters out some species and favors others on the basis of specific species traits. For example, sites on xeric, exposed slopes generally contain more drought tolerant plants than do mesic sites. The species traits may be phenotypic (e.g., drought tolerance), but they may also be biogeographic (e.g., native vs. nonnative), taxonomic, or ecological (e.g., trophic position). Even when beta diversity or competition are high and few sites share species, we expect the *traits* expressed at sites to be consistent with their environments.

Just as individuals represent alleles in a population, species represent their traits within a community. For n species occupying m sites and possessing a alternative levels of a categorical trait (hereafter *states*), we can construct an $m \times n$ matrix, $\mathbf{S}$, of species abundances (or occurrence, biomass, importance, etc.) at sites and an $n \times a$ matrix, $\mathbf{W}$, of state expression by species. The columns in $\mathbf{W}$ represent the alternative states associated with a single trait, rather than several separate traits as in RLQ analysis. Where species may show more than one state, the values in $\mathbf{W}$ correspond to the probability that any given individual of species k will express state j (Chevenet et al. 1994). For instance, if the seeds of a certain tree species are usually dispersed by the wind but are occasionally dispersed by birds, the species might receive a score of 0.8 for wind dispersal and a score of 0.2 for bird dispersal. Because the rows in **W** represent probability density functions for trait expression by species, their values must be nonnegative and must sum to one.

To calculate the trait selection ratios, we must (1) translate species abundances at sites into CAT scores for the sites, and (2) quantify the availability of each state to each site. For a categorical trait, the expression of state j at site i can be estimated as the proportional abundance of species that posses j and occur at i, $\hat{p}_{ij}$. The set $\hat{p}_{ij} : (\hat{p}_{i1}, \ldots, \hat{p}_{ia})$ represents the proportional abundances for the alternative states at i; it is the equivalent to the CAT score for a discrete trait. To calculate $\hat{p}_{ij}$, we can sum the species abundances at i, weighting each species by its probability of expressing j:

$$\mathbf{G} = \mathbf{SW}. \qquad (3)$$

The rows of the grouped matrix, **G**, are then standardized by the total abundance of all species at site i (i.e., the row sums of **S**).

The availability of state j at site i (π_{ij}) equals the pooled, proportional abundance of the state's carrier species in the site's regional species pool, i.e., $\mathbf{S}'\mathbf{W}$, where $\mathbf{S}'$ is a matrix of proportional species abundances in the site-specific species pools. The specification of $\mathbf{S}'$ is critical and depends on a null model appropriate to the specific system being analyzed (see *Discussion*). The null model can be as complex as necessary for the study system. For the sake of simplicity, we assume here that all sites draw from the same pool. In this case, $\hat{\pi}_{ij}$ can be estimated by standardizing the column sums of **S** to the grand total of **S** and multiplying the resulting vector by **W**. We refer to this estimation as the site pooling model. Ideally, the values of $\hat{\pi}_{ij}$ should be estimated separately from the data used to calculate $\hat{p}_{ij}$; availabilities calculated by site pooling will almost certainly contain biases. However, site pooling may be useful if **S** contains the only available data on the study system and the assumption of a single species pool is not unreasonable. Once $\hat{p}_{ij}$ and $\hat{\pi}_{ij}$ have been calculated, Eqs. 1 and 2 estimate the filters at site i as the probability that i selects for state j, $\hat{B}_{ij}$.

Selection ratios may also be calculated for continuous traits. Using the definition of matrix **W** as a set of conditional probability functions for trait expression by species, the CAT score of Garnier et al. (2004) can be obtained by assigning a density of 1 to each species' mean trait score in **W** and then calculating the row means of **G**. In general, if the kth row in **W** represents the probability $f_k(t)$ that species k expresses trait value t, then **G** becomes a mixture model for the community aggregated trait distribution at site i, $f_i(t)$:

$$\hat{p}_i \propto \mathbf{G}_i = f_i(t) = \sum_{k=1}^{n} \left(\frac{\mathbf{S}_{ik}}{\mathbf{S}_{i+}}\right) f_k(t). \qquad (4)$$

This formulation allows the analysis to make use of whatever information is available on trait expression by species, with the caveat that a poorly defined $f_k(t)$ will yield less precise (and possibly biased) estimates of $f_i(t)$ (see Appendix). As in the discrete case, the denominator of the selection ratio is found by recalculating **G** with relative abundances expected under an appropriate null model of species availability.

Selection (sel) and availability (avl) are assumed to be independent and to be the only processes that affect the probability of observing (obs) an aggregated trait value of t at i (i.e., $\Pr[t_{\mathrm{obs}}] = \Pr[t_{\mathrm{avl}}] \times \Pr[t_{\mathrm{sel}}]$). If availability has been correctly described by the null model, the selection ratio for $f_i(t)$ therefore estimates the selective filter in operation at site i as a probability distribution on t. Thus, $f_i(t)$ is the equivalent of $\hat{B}_{ij}$ for a continuous trait. In practice, $f_i(t)$ can be approximated by dividing the range of t into a large number of discrete bins and calculating $\hat{B}_{ij}$ (Appendix). This approach is nonparametric. If $f_i(t)$ is assumed to follow a specific distribution, parameters can be estimated by applying the method of moments to a bootstrap sample of t following $\hat{B}_{ij}$.

Differences among the estimated filters at different sites do not necessarily indicate selection, as some variation must occur due to random sampling. To test for significant deviations from the null model, a predictive envelope for $\hat{B}_{ij}$ can be obtained by estimating a bootstrap sample of the filter, $\hat{B}^{*}_{ij}$, from several ($\geq$200) simulated samples of community composition under the null model. To do this, we fix the total abundance at each site (row sums in **S**) and randomly assign individuals to species by inverse distribution sampling from the species pool proportions (column sums in $\mathbf{S}'$). Under the site pooling model, species membership at site i is governed by the following function:

$$F_i(k \leq m) = \sum_{k=1}^{m} \frac{\mathbf{S}_{+k}}{\mathbf{S}_{++}} \qquad (5)$$

where $F_i(k)$ is the probability of encountering the kth species in the pool at site i and $F_i(0) = 0$. To assign a species to an individual, a uniform random number, U (between 0 and 1), is compared to $F_i(k)$, and k is selected so that $F_i(k-1) < U \leq F_i(k)$. Repeating the process for $\mathbf{S}_{i+}$ individuals produces a simulated sample from the community at i under the null model. There is evidence for significant selection by an individual site if the difference between the observed filter ($\hat{B}_{ij}$) and the bootstrap filters ($\hat{B}^{*}_{ij}$) is significantly greater than the pairwise differences among the bootstrap replicates. This can be tested by a bootstrap Kolmogorov-Smirnov test (Appendix). In general, the null model can be rejected if a significant number of sites (i.e., >5%) show selection. In this case, the filter means can be modeled by regression. If filters are assumed to take a non-normal parametric form, a GLM approach may be more appropriate. Alternatively, quantile regression might

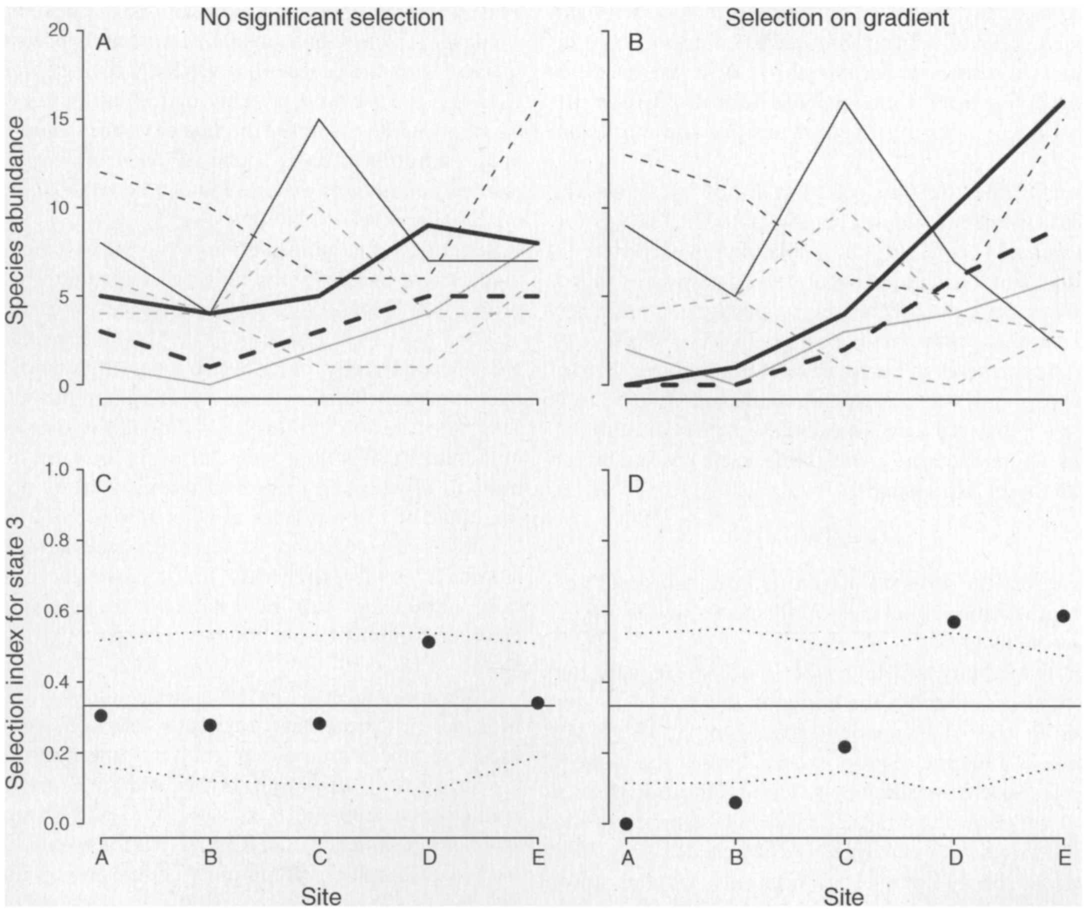

FIG. 1. (A, B) Species abundances and (C, D) standardized selection ratios at five sites (simulated data). Seven species (different line styles and shades in panels A and B) each express one of three trait states (gray, black, and bold lines in upper panels). In the lower panels, points indicate the standardized selection index, $\hat{B}$, for the third state at each site. Dashed and dotted lines indicate the expected value (=1/3) and the Bonferroni-corrected 95% bootstrap prediction interval for the null model of no selection, respectively ($n = 200$ replicates). Marginal species abundances are equal for both cases.

be used for a completely non-parametric approach (Cade and Noon 2003). R code to calculate selection ratios and to perform bootstrap tests is available in Supplement 1.

EXAMPLES

A numerical example

This example uses a hypothetical data set with abundances of seven species at five sites. Each species exhibits one of three states of a discrete trait. We first consider a case in which no selection occurs (Fig. 1A), and then a case in which sites differ in selection for state 3, but not for states 1 and 2. In both cases, species have known expected proportions (π_{ij}) of 0.8, 0.13, 0.09, 0.21, 0.26, 0.15, and 0.08, respectively. Data and calculations for this example can be found in the Appendix, along with a similar example for a continuous trait.

A chi-squared test (Manly et al. 2002:65; Appendix) on **S** reveals that sites differ significantly in species composition (Manly's $\chi_1^2 = 38.98$, df = 24, $P < 0.001$), but a similar test performed on **G** shows no differences among sites for trait selection ($\chi_1^2 = 5.62$, df = 8, $P = 0.696$). This is confirmed by plotting the individual site filters (i.e., $\hat{B}_{ij}$) and prediction intervals for $\hat{B}^*_{ij}$ (Fig. 1C). All of the standardized selection indices fall within the 95% confidence band.

In the second case, the marginal values for **S** remain unchanged, but species composition is rearranged so that expression of state 3 responds to a gradient running from Site A to Site E (compare Fig. 1A, B). On average, states occur in proportion to their regional availability (Manly's $\chi_3^2 = 0.29$, df = 2, $P = 0.864$), but individual sites differ significantly in their selection for or against state 3 ($\chi_1^2 = 51.25$, df = 8, $P < 0.001$; Fig. 1D). Specifically, sites A and B select against state 3, while D and E select for it. Subsequent analysis would attempt to relate the observed increase in selection from A to E to changes in environmental conditions across the five sites.

Example 2: selection for nonnative plants

Our second example illustrates the use of standardized selection ratios and linear regression to relate selection

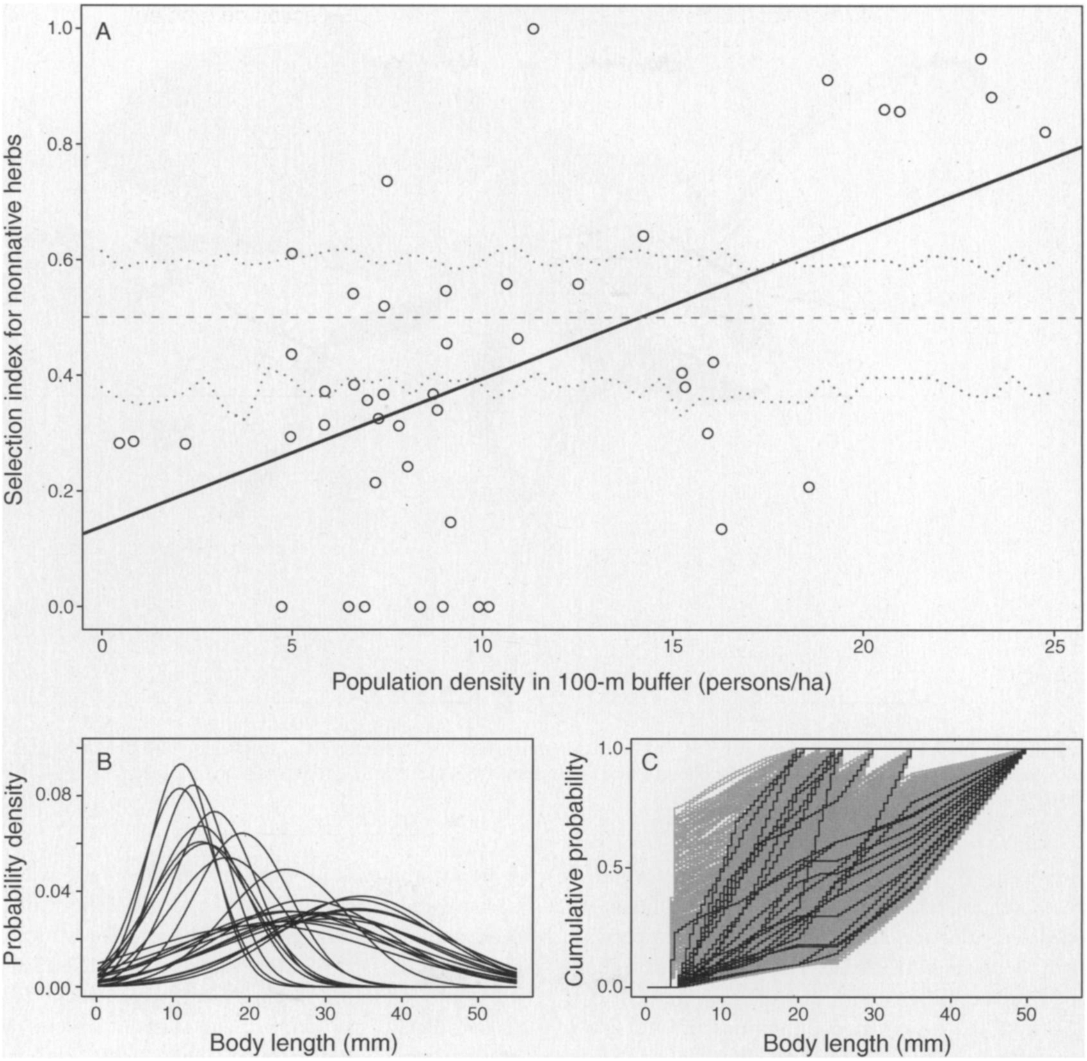

FIG. 2. Examples of the use of selection ratios on community aggregated traits to estimate ecological filters. (A) Greater human population density adjacent to urban natural areas increases selection for nonnative herbs in Portland, Oregon ($\hat{B} = 0.138 + 0.026 \times$ POPDENS; $R^2 = 0.301$, $F_{1,45} = 19.33$, $P < 0.001$, bold line). Points show estimates of $\hat{B}$ for nonnative vs. native herbaceous plants, relative to population density in a 100-m buffer around study sites. A value of 0.5 (dashed line) indicates neutral selection. Values outside the 95% bootstrap confidence envelope (dotted lines) indicate significant selection for (above the envelope) or against (below) nonnative species. Bootstrap intervals were based on 200 independent random communities simulated from the empirical distribution function of species abundance pooled across sites, assuming no selection. (B, C) Estimated selection functions for body size in a community of 81 Cerambycid beetle species appears to vary among 23 forest sites in Indiana (panel B; curves show estimates for individual sites assuming normal filters), but the observed set of filters (panel C, black lines) are consistent with the set produced by the null model (C, gray lines, $n = 200$ bootstrap replicates). In bootstrap Kolmogorov-Smirnov tests, the difference between null model and observed filters was greater than the difference between pairs of null model realizations for only one site (bootstrap $P = 0.037$, $n = 199$ null pairs; all other sites $P > 0.24$).

for a specific state to an environmental variable. We ask how selection for nonnative vs. native herbs relates to adjacent human population density in 47 green spaces in the Portland, Oregon, USA metropolitan area. Herbs were sampled during the summers of 2001–2002 and species importance values were calculated for each site. We estimated the human population density in a 100-m buffer around each site using 2000 U.S. Census tract data in ArcGIS 8.2 (ESRI, Redlands, California, USA).

Species were classified as native or nonnative to western Oregon, and selection ratios, confidence intervals, and Manly's first chi-squared statistic were estimated using the functions in Supplement 1, run in R version 2.6.2 (R Development Core Team 2007). The data for this example appear in Supplement 2. We use the site pooling model to define availability. Because this case involves only two states, their standardized selection indices are complementary (i.e., $\hat{B}(\text{native}) = 1$

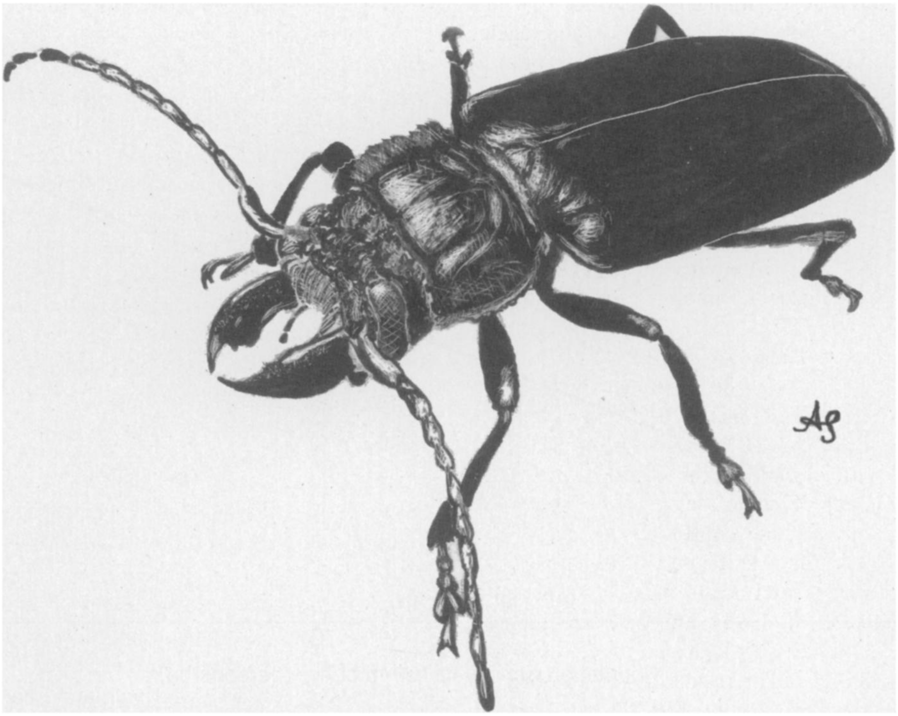

PLATE 1. *Mallodon dasystomus* (Say), an example of a longhorn beetle (Coleoptera: Cerambycidae) from Mississippi, USA. Variation in body size among longhorn beetle communities in Indiana (USA) forest fragments is analyzed (see *Examples: Example 3*). Illustration: Annie Elizabeth Spikes.

– $\hat{B}$(non-native)). This would not be the case for a more complex set of characteristics. Fig. 2 shows the standardized selection indices for nonnative species in relation to population density, together with a bootstrapped prediction envelope for the null model. More than 63% of the observed communities fall outside of this envelope, clearly indicating that selection has occurred. Linear regression on population density explains 30% of the variation in $\hat{B}$ among sites ($\hat{B} = 0.138 + 0.026 \times \text{POPDENS}$, $R^2 = 0.301$, $F_{1,45} = 19.33$, $P < 0.001$); a more appropriate model that treats $\hat{B}$ as a beta random variable would improve the fit considerably. Interestingly, selection is not reflected by the richness or diversity of nonnative herbs in this case. Regressions of these indices produced R^2 values of only ~0.5% (richness, $F_{1,45} = 2.184$, $P = 0.146$; Shannon's H, $F_{1,45} = 1.444$, $P = 0.236$; Simpson's D, $F_{1,45} = 0.233$, $P = 0.622$).

Example 3: a continuous trait with no evidence of selection

For this example, we use data on longhorned beetle (Coleoptera: Cerambycidae) (see Plate 1) assemblages at 23 study sites in Indiana, USA. Depending on species, longhorned beetles lay their eggs in live or dead wood; larvae feed on the host log or tree. After pupating, the adult beetles disperse in search of mates and appropriate oviposition sites of their own. We hypothesized that if dispersal ability is proportional to body size within the Cerambycidae (measured as length), selection for larger beetles might correlate negatively with forest density in the landscape surrounding capture sites. The data set, sites, and sampling methods are presented in Holland (2006). Beetle size data, in the form of species' length ranges, were obtained from Yanega (1996). This analysis follows the procedure for continuous traits outlined in the Appendix, and uses the R functions from Supplement 1. The estimated filter functions vary widely among sites, with some sites apparently favoring small beetles while others favored larger ones (Fig. 2B). However, the observed pattern of filters fell entirely within the cloud generated by bootstrap sampling under the site pooling model (Fig. 2C, bootstrap Kolmogorov-Smirnov P values >0.24 for 22 of 23 sites, $n = 199$ null model replicates). We therefore failed to reject the hypothesis that the observed distributions of beetle sizes were derived by random sampling from the regional community.

DISCUSSION

As a solution to the central problem of fourth-corner analyses, selection ratios on community aggregated traits (SRCAT) are easy to calculate and may be adapted to accommodate different levels of information on species traits. They also provide a useful conceptual link to the theory of ecological filters. Most existing fourth-corner approaches implicitly assume that observed trait–environment correlations or aggregated trait values do not arise from confounding spatial or temporal differences in the species' availability to colonize sites. However, unequal propagule pressure is often the rule rather than the exception. SRCAT requires that assumptions about species availability be explicitly described in a null model, and then accounts for differential availability in the filter estimates. Because the filters are probability distributions, they are also directly comparable among different regional communities (Manly 1993).

RLQ analysis (Dolédec et al. 1996, Dray and Legendre 2008), of which the original fourth-corner analysis (Legendre et al. 1997) is a special case, possesses distinct advantages over SRCAT for some applications. RLQ simultaneously estimates the correlations among multiple environmental covariates and species traits, each of which may include continuous or categorical variables (or both). Its results may be visualized as ordinations of species and traits onto environmental gradients (Dolédec et al. 1996), and permutation tests are available for various ecological hypotheses (Dray and Legendre 2008). Similar results might be obtained through a two-step process of SRCAT analysis on several traits followed by constrained ordination, but this approach decouples species from the second stage of the analysis. On the other hand, site-level SRCAT functions can be analyzed by standard statistical approaches, and therefore provide greater flexibility from a modeling or prediction perspective. Ultimately, it would be advantageous to combine these approaches into a single analysis that allows simultaneous, generalized modeling while accounting for variation in species availability among sites.

Both RLQ and SRCAT can incorporate intraspecific variation through fuzzy coding. This practice assumes that variation within species is not related to environmental differences among sites, which is rarely true. Environmentally induced trait plasticity is relatively common, so intraspecific trait expression should vary systematically among sites if strong filters exist (Agrawal 2001). To obtain the best possible filter estimates, the rows in **S** and columns in **W** should ideally represent individual populations, rather than species, and the trait should be measured at each site. As the quality of trait data declines from population-level to species-level distributions, and finally to species-level trait ranges or means, the power to detect significant filters is also expected to decline. In general, we expect that analyses based on species-level trait data will overestimate the variance in site's ecological filters.

Filter estimation by SRCAT depends heavily on the null model that defines the sites' regional species pools. As with any null model approach, practitioners should be careful to select an appropriate model for their study system and questions (Harvey et al. 1983). Studies that cover large geographic regions, multiple habitat types, or long time spans will generally require more complex null models to account for confounding species-pool effects, unless the factors that determine availability (e.g., distance from source populations) are themselves treated as environmental covariates driving selection. Where uncertainty exists, it may be beneficial to compare results under a range of models for availability (Kelt et al. 1995). In some cases, it also may be sufficient to know the direction in which any biases lie. For example, the estimated availability of exotic plants in our second example is probably too low, since these species are more abundant in the urban matrix than in the sampled reserves. Any apparent selection against non-natives therefore represents a worst-case scenario with regard to exotic invasion.

CONCLUSIONS

Shipley and colleagues (2006) demonstrate that analyses based on CAT scores can use data on species' traits and environmental covariates to correctly predict species' site-specific relative abundances. Our generalization of the CAT index and addition of a null model for species availability to sites increases the flexibility of this approach by allowing for qualitative traits and by relaxing two main assumptions: that intraspecific variation is negligible compared to differences among species (Garnier et al. 2004, Shipley et al. 2006), and that sites draw colonists from a panmictic community. Modeling changes in the filter function rather than the CAT value should improve predictive power and allows analysis of extreme values as well as mean trends (Cade and Noon 2003). In addition, we hope that generalization to discrete variables will also facilitate analysis of continuous traits. Community ecologists often collect data on species composition and environmental variables, but only rarely measure species traits as part of community-level surveys. This approach allows researchers to draw trait information from the literature or from independent databases where quantitative traits may appear as ordinal values.

ACKNOWLEDGMENTS

Critiques and suggestions from Robert Swihart, Jeffery Holland, Andrew Muir, and two anonymous reviewers greatly improved the current paper. We also thank Ruth Hufbauer and two anonymous reviewers for their comments on a previous manuscript describing the trait-selection method, and J. Holland, for providing the beetle data used in example 3.

LITERATURE CITED

Agrawal, A. A. 2001. Phenotypic plasticity in the interactions and evolution of species. Science 294:321–326.

Cade, B. S., and B. R. Noon. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1:412–420.

Chevenet, F., S. Doledec, and D. Chessel. 1994. A fuzzy coding approach for the analysis of long-term ecological data. Freshwater Biology 31:295–309.

Darwin, C. 1859. On the origin of species. Murray, London, UK.

Díaz, S., S. Lavorel, F. S. Chapin, III, P. A. Tecco, D. E. Gurvich, and K. Grigulis. 2007. Functional diversity—at the crossroads between ecosystem functioning and environmental filters. Pages 81–91 *in* J. G. Canadell, D. E. Pataki, and L. F. Pitelka, editors. Terrestrial ecosystems in a changing world. Springer, New York, New York, USA.

Dolédec, S., D. Chessel, C. J. F. ter Braak, and S. Champely. 1996. Matching species traits to environmental variables: a new three-table ordination method. Environmental and Ecological Statistics 3:143–166.

Dray, S., and P. Legendre. 2008. Testing the species traits–environment relationships: the fourth-corner problem revisited. Ecology 89:3400–3412.

Garnier, E., J. Cortez, G. Billes, M. L. Navas, C. Roumet, M. Debussche, G. Laurent, A. Blanchard, D. Aubry, and A. Bellmann. 2004. Plant functional markers capture ecosystem properties during secondary succession. Ecology 85:2630–2637.

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

Holland, J. D. 2006. Cerambycidae larval host condition predicts trap efficiency. Environmental Entomology 35: 1647–1653.

Hutchinson, G. E. 1957. Concluding remarks. Population studies: animal ecology and demography. Cold Spring Harbor Symposia on Quantitative Biology 22:415–427.

Kelt, D. A., M. L. Taper, and P. L. Meserve. 1995. Assessing the impact of competition on community assembly: a case study using small mammals. Ecology 76:1283–1296.

Legendre, P., R. Galzin, and M. L. Harmelin-Vivien. 1997. Relating behavior to habitat: solutions to the fourth-corner problem. Ecology 78:547–562.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Second English edition. Elsevier Science, Amsterdam, The Netherlands.

Manly, B. F. J. 1993. Comments on design and analysis of multiple-choice feeding-preference experiments. Oecologia 93:149–152.

Manly, B. F. J., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: statistical design and analysis for field studies. Second edition. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Poff, N. L. R. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391–409.

R Development Core Team. 2007. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org⟩

Savage, R. E. 1931. The relation between the feeding of the herring off the east coast of England and the plankton of the surrounding waters. Fishery Investigation, Ministry of Agriculture, Food, and Fisheries, Series 2 12:1–88.

Shipley, B., D. Vile, and E. Garnier. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 314:812–814.

Strickland, M. D., and L. L. McDonald. 2006. Introduction to the special section on resource selection. Journal of Wildlife Management 70:321–323.

Weiher, E., and P. A. Keddy. 1995. The assembly of experimental wetland plant communities. Oikos 73:323–335.

Yanega, D. 1996. Field guide to northeastern longhorned beetles (Coleoptera: Cerambycidae). Illinois Natural History Survey, Champaign, Illinois, USA.

APPENDIX

Estimating ecological filters with selection ratios and community aggregated traits (*Ecological Archives* E091-026-A1).

SUPPLEMENT 1

R source code for functions and examples of selection ratio analysis on community aggregated traits (SRCAT) (*Ecological Archives* E091-026-S1).

SUPPLEMENT 2

Data for examples of selection ratio analysis on community aggregated traits (SRCAT) (*Ecological Archives* E091-026-S2).