

A power analysis for multivariate tests of temporal trend in species composition

KATHRYN M. IRVINE,^{1,3} ERIC C. DINGER,² AND DANIEL SARR²

¹*Department of Mathematical Sciences, Montana State University, Bozeman, Montana 59717 USA*

²*National Park Service Klamath Network, 1250 Siskiyou Boulevard, Ashland, Oregon 97520 USA*

Abstract. Long-term monitoring programs emphasize power analysis as a tool to determine the sampling effort necessary to effectively document ecologically significant changes in ecosystems. Programs that monitor entire multispecies assemblages require a method for determining the power of multivariate statistical models to detect trend. We provide a method to simulate presence–absence species assemblage data that are consistent with increasing or decreasing directional change in species composition within multiple sites. This step is the foundation for using Monte Carlo methods to approximate the power of any multivariate method for detecting temporal trends. We focus on comparing the power of the Mantel test, permutational multivariate analysis of variance, and constrained analysis of principal coordinates. We find that the power of the various methods we investigate is sensitive to the number of species in the community, univariate species patterns, and the number of sites sampled over time. For increasing directional change scenarios, constrained analysis of principal coordinates was as or more powerful than permutational multivariate analysis of variance, the Mantel test was the least powerful. However, in our investigation of decreasing directional change, the Mantel test was typically as or more powerful than the other models.

Key words: *Bray-Curtis; constrained analysis of principle coordinates; long-term monitoring; Mantel test; PerMANOVA; presence–absence data; vegan R package.*

INTRODUCTION

Two critical questions in designing environmental monitoring programs are (1) What to monitor? and (2) How do we rigorously document important trends in the ecological indicators selected? Funding limitations demand that programs must be selective in what they monitor; historically, populations of single species have been emphasized (e.g., grizzly bears; Eberhardt et al. 1986). Alternatively, biological assemblages can be monitored, such as aquatic macroinvertebrates (e.g., Rosenberg and Resh 1993) or vertebrates (e.g., Manley et al. 2004). Monitoring programs wisely require documenting the level of change that can be detected prior to implementation, as in the Oakley et al. (2003) guidelines originally written for the National Park Service, now used by many other monitoring entities. Typically, in the analysis of multispecies assemblage data, the species response matrix is transformed into a matrix of pairwise dissimilarity values prior to analysis. However, evaluating power to detect trends in such data remains problematic (Sommerfield et al. 2002). We present a novel way to explore and estimate the power of dissimilarity-based multivariate analyses to inform the

allocation of effort (sample size) in long-term monitoring programs focusing on multispecies assemblages.

Previous multivariate power studies have focused on hypothesis tests between a priori groups (Sommerfield et al. 2002, Anderson and Robinson 2003, Warton and Hudson 2004) and those exploring temporal trends assumed the underlying species distributions were continuous random variables and used the Euclidean distance measure (Collins et al. 2000, Angeler et al. 2009; although see Legendre et al. 2005 for binary data in a spatial setting). Here we consider the case of binary species responses and the Bray-Curtis measure of dissimilarity for presence/absence data. Others have mentioned that the choice of distance measure can affect the interpretation of patterns in the data (Anderson and Robinson 2003, Warton and Hudson 2004). Accordingly, we provide insights into the interpretation of temporal trends in assemblages summarized based on the Bray-Curtis measure for binary data.

Temporal trends in community data summarized in terms of pairwise dissimilarities in species composition can be considered in multiple ways (Philippi et al. 1998, Collins et al. 2000). If the same sites are always monitored over time, one can consider linear trends in the average dissimilarity values in which the average is calculated using pairs of observations within a year (among sites) or within a site (among years). In terms of *among* site temporal trend in species composition we distinguish two possible patterns: convergence or

Manuscript received 5 November 2010; revised 25 May 2011; accepted 6 June 2011. Corresponding Editor: N. J. Gotelli.

³ Present address: U S Geological Survey, Northern Rocky Mountain Science Center, 2327 University Way, Suite 2, Bozeman, Montana 59715 USA; E-mail: kirvine@usgs.gov

divergence. Commonly, convergence in species composition means the average dissimilarity *among* sites within a year decreases over time, increasing assemblage homogeneity across a landscape. Whereas divergence in species composition means the average dissimilarity *among* sites within a year is increasing over time, the species assemblages are becoming more distinct from each other. In terms of temporal trends in species composition *within* a site, we focus on two possible patterns: increasing or decreasing directional change. We define directional change as linearly increasing or decreasing average dissimilarity in species composition *within* a site over increasing time lags.

Ecologically, increasing directional change can occur with a pressing disturbance (e.g., climate change or atmospheric deposition), where the stressor is constantly impacting the assemblage and furthering species turnover within a site over time. On the other hand, decreasing directional change is a pattern resulting from a short-term disturbance. For example, consider the theoretical compositional shift of aquatic macroinvertebrates along a gradient of increasing sedimentation over time at only one site. A low sedimentation macroinvertebrate community is largely mayflies, stoneflies, and caddisflies, but the community shifts toward being dominated by worms, true flies, and other more tolerant taxa as sedimentation increases; three groups of species are decreasing and three are increasing over time. If, within the time frame of monitoring sedimentation returns to low levels, the species composition should revert to the macroinvertebrate community observed initially (a short-term disturbance). We assume there is a degree of temporal coherence among sites meaning all sites are affected by the same pressing or short term disturbance. Here, we focus on within site trends in species composition because a long term monitoring program would be interested in verifying the specified sampling effort of replicated sites yields enough power to detect either increasing or decreasing directional change in the target population of all possible sites.

The major complication is that the species within any given community type can change in an effectively infinite number of ways, and, presently, it is not clear which patterns generate communities consistent with increasing or decreasing directional change. Consequently, here, we provide a method for generating binary species assemblage data that display directional change in species composition within multiple sites over time (see Supplement 2 for R code). This step is the foundation for using Monte Carlo methods to approximate the power of any multivariate statistical method to detect temporal trend in species composition. Our methodology provides a robust method for calculating the needed sample size to achieve a specified level of power, a necessary step to justify the cost of long-term monitoring of multi-species assemblages.

In this paper, we consider three different multivariate tests for trend: (1) the Mantel test (Philippi et al. 1998,

Somerfield et al. 2002); (2) permutational multivariate analysis of variance (PerMANOVA; McArdle and Anderson 2001); and (3) constrained analysis of principle coordinates (CAP; Anderson and Willis 2003). We use the R 2.7.2 statistical software package (R Development Core Team 2008) for a simulation investigation of the power of these different models to detect increasing or decreasing directional change. We also explore the influence of the underlying univariate species patterns, number of species in the community, and number of sites sampled over time on the power of these dissimilarity-based models.

METHODS

Statistical analyses

The three methods for detecting trends in species composition we explored (Mantel, CAP, and PerMANOVA) are easily implemented within the *vegan* package in R (see Supplement 1 for details). We investigated the situation in which n sites are randomly selected for monitoring and all sites are visited every survey period t for a total of $N (= nT)$ unique rows in a data matrix (\mathbf{Y}). The number of columns in \mathbf{Y} , denoted as p , depends on the species list used; e.g., macroinvertebrate species collected in one or more kick net samples. We focused on the case of species assemblage data gathered as simply presence or absence, e.g., the elements of \mathbf{Y} are either 0 or 1. We used the Bray-Curtis dissimilarity measure for binary data, $d_{ij} = [(\sum_k |y_{kj} - y_{ki}|) / (\sum_k (y_{kj} + y_{ki}))]$, where j and i index the rows and k indexes the columns of \mathbf{Y} , the $N \times N$ distance matrix \mathbf{D} is composed of elements d_{ij} .

The Mantel test has a long history in ecological applications for relating two distance matrices (e.g., Legendre and Legendre 1998). In our setting, the Spearman correlation based on ranks was calculated between the upper diagonal elements in matrix \mathbf{D} and a matrix of temporal lags. The temporal lag matrix was found by calculating the Euclidean distance measure on a vector composed of an integer value for the year of sampling t (*vegdist* in the *vegan* package of R; Oksanen et al. 2010). Since all sites are visited every year we used a constrained permutation to calculate the significance of the Spearman correlation statistic. We used an augmented version of the Mantel function in the R package *vegan* because we were interested in both evidence of an increase or decrease in dissimilarity with an increase in time lag (example in Supplement 1).

McArdle and Anderson (2001) popularized the use of multi-response permutation methods for community analyses. Essentially, any distance matrix (\mathbf{D}) can be partitioned into a linear model. In our case, the design matrix \mathbf{X} contains a column of ones, $n - 1$ indicators for the sites, and a column composed of integer values representing the year of sampling (e.g., 0, 1, 2). The sequential sum of squares F statistic based on a model with and without the year column was the test statistic for trend. A permutation P value was calculated to

assess statistical significance. We used a constrained permutation to account for the fact the same sites were visited every year by permuting the temporal observations within a site. We used the *adonis* function in the *vegan* package of R to fit this model for multivariate trend (e.g., Supplement 1; Oksanen et al. 2010).

Canonical analysis (or constrained analysis) of principal coordinates analysis can be used on any distance measure and is equivalent to redundancy analysis when the Euclidean distance is used (e.g., Anderson and Willis 2003). The test statistic in CAP was based on the sum of squared canonical correlations (for mathematical details, see Anderson and Willis 2003). A benefit of CAP is an ordination can be produced which displays the site temporal trajectories in the p -dimensional species space. This type of ordination could reveal trends not present within a nonmetric multidimensional scaling of the same data. We used the *capscale* function in the R package *vegan* to implement this test (example in Supplement 1; Oksanen et al. 2010).

Simulating assemblage data consistent with directional change

We provide R code to simulate presence/absence community data that display either an increase or decrease in average dissimilarity in species composition within a site over increasing time lags for multiple sites (Supplement 2). We assume that species within the same group have a common average deterministic trend over time. Also, we assume a degree of temporal coherence among sites; the ecological processes effecting species composition within a site are exerting an influence across all sites (Larsen et al. 2001); i.e., all sites display increasing or decreasing directional change. Our methods are appropriate for programs that use a fixed species list and monitor permanent sites over time (always revisit design).

Specifically, the observation at a site for one species is a binary random variable y_{ikt} where i indexes the site, $i = 1, \dots, n$; k indexes the species, $k = 1, \dots, p$; and t indexes the year of sampling, $t = 1, \dots, T$. The probability of a species being present is ϕ_{ikt} (occurrence probability), so $y_{ikt} | \phi_{ikt} \sim \text{Bernoulli}(\phi_{ikt})$. Assuming perfect detectability, a logistic regression model with quadratic trend for each species is $\text{logit}(\phi_{ikt}) = a_{0i} + b_{0k} + a_{1i}Yr_t + b_{1k}Yr_t + b_{2k}Yr_t^2$, where ϕ_{ikt} is the probability of occurrence for species k at site i at time t , Yr_t is the real number assigned to year t of sampling. Notice in this model if we assume $b_{2k} = 0$ for all k species, then each unique species and site combination has a different trend line with intercept $a_{0i} + b_{0k}$ and slope $a_{1i} + b_{1k}$. The full model assumes each species has a common quadratic trend across sites; and each site has a common trend line a_{0i} intercept and a_{1i} slope. Notice we are modeling trend in terms of a deterministic function of year.

Assume, for example, that there are three groups of species ($K = 1, 2, 3$) where g_K is the number of species

within each group ($\sum_K g_K = p$). To incorporate different group effects we specify the following higher level:

$$\begin{Bmatrix} b_{0k} \\ b_{1k} \\ b_{2k} \end{Bmatrix} \sim \text{MVN}_3 \left\{ \begin{Bmatrix} \mu_{0K} \\ \mu_{1K} \\ \mu_{2K} \end{Bmatrix}, \begin{Bmatrix} \sigma_{0K}^2 & \sigma_{01K}^2 & \sigma_{02K}^2 \\ \sigma_{01K}^2 & \sigma_{1K}^2 & \sigma_{12K}^2 \\ \sigma_{02K}^2 & \sigma_{12K}^2 & \sigma_{2K}^2 \end{Bmatrix} \right\}$$

if k is in group K . This model is different from a mixed model with random species effect (random intercepts and slopes for each species) because we allow for “group” effects, such that species in different groups have different patterns, but we are inducing correlation among species within the same group. We assume permanent sites for monitoring. To model the repeated measures over time we specify a common intercept and slope for all species within site i by adding another level as follows:

$$\begin{Bmatrix} a_{0i} \\ a_{1i} \end{Bmatrix} \sim \text{MVN}_2 \left\{ \begin{Bmatrix} \tau_0 \\ \tau_1 \end{Bmatrix}, \begin{Bmatrix} \gamma_0^2 & \gamma_{01}^2 \\ \gamma_{01}^2 & \gamma_1^2 \end{Bmatrix} \right\}.$$

Notice the variance parameters ($\gamma_0^2, \gamma_{01}^2, \gamma_1^2$ influence the degree of temporal coherence among sites or site-to-site variability.

The Bray-Curtis measure of dissimilarity for binary data is simply the number of unique species in rows j and i divided by sum of the number of species in row j and the number of species in row i . For average dissimilarity to increase within a site over increasing time lags (increasing directional change), the number of unique species must increase over time indicating a high species turnover within a community, assuming a fixed species pool. On the other hand, for average dissimilarity to decrease within a site over increasing time lag (decreasing directional change), the number of unique species must first increase (say, following disturbance), and then decrease (say, after cessation of the disturbance) during the period relative to a starting baseline composition. For example, species observed in year t of sampling are more similar to the suite of species observed in, say, year $t + 10$ of sampling than the intervening years (e.g., year t to year $t + 1$). Consequently, to simulate a species assemblage within a site undergoing decreasing directional change we assume that ϕ_{ikt} is a quadratic function of year for a species k and Yr_t is mean centered, $Yr_t = t - \bar{t}$. Whereas increasing directional change assumes each univariate species pattern is a linear function of year, $b_{2k} = 0$ for all k and $Yr_t = t - 1$. To simulate an assemblage without directional change (no deterministic temporal drivers effecting species composition within a site) we set all mean and variance parameters with subscript 1 or 2 to zero; e.g., $\mu_{1K} = \mu_{2K} = \tau_1 = \gamma_1^2 = 0$ (for an alternative null model approach for a single site, see Gotelli et al. [2010]). These methods allow us to simulate presence-absence data conforming to the ecological conjecture of directional change in species composition within multiple sites monitored over time. In Appendix A, we provide examples of simulated communities undergoing increasing and decreasing directional change in species

composition. Also, we show the resulting patterns in dissimilarity among sites within time periods.

For this simulation method to be linked to realistic variability among sites and species, the variance parameters must be estimated from available pilot data or gathered from the available literature. If pilot data are available, fitting separate logistic regression models to each of the more ubiquitous species (no pooling) or a mixed model with random species effects (partial pooling) can be used to find reasonable values for the underlying variance parameters. For our investigation, we used a long-term data set available from the PACFISH/INFISH biological opinion monitoring program consisting of 84 aquatic macroinvertebrates species sampled at eight sites for eight years (Kershner et al. 2004). This particular dataset, although measured as abundance counts, was reduced to binary data. There was very little evidence of site-to-site and species-to-species variation within this particular dataset.

Power analysis

We use a Monte Carlo approach to estimate the power of the Mantel test, PerMANOVA, and CAP for detecting directional change in species assemblages. We generate multivariate presence/absence data for a given alternative hypothesis of interest (increasing, decreasing, and no directional change), variance parameters, number of sites (n), and number of years (T) using the method previously described. For each iteration of simulated data, each model was fit (Mantel, PerMANOVA, CAP), a permutation-based P value was calculated for trend in each model, and the number of P values $\leq \alpha$ level were counted and divided by the total number of iterations (in our case 100) for each model, yielding an estimate of power for each model under the specified conditions. We set $\alpha = 0.05$. The specific parameter values used in the simulation study are provided in Appendix B.

We investigate the power assuming 10, 20, and 40 sites were sampled for 10 years (e.g., $nT = 100, 200, \text{ or } 400$). Also, we vary the number of total species ($p = 40$ or 100). We are primarily interested in the sensitivity to the underlying univariate species temporal patterns. For the increasing directional change scenarios, we assume the same magnitude of average annual multiplicative change in the odds for a species group (0% or 5% annual multiplicative change in the odds, on average). We vary the sign of the parameter such that within a group all species display either a multiplicative increase or decrease in their odds of occurrence over time, on average. For increasing directional change, we explore three different species patterns: (1) 50% of the species increase and 50% with no trend, (2) 50% of the species decrease and 50% no trend; and (3) 50% with increasing and 50% decreasing (high species turnover). For decreasing directional change, we alter the number of species showing concave up trends over time (90%, 50%, 10% with $\mu_{1K} = 0.05, \mu_{2K} = 0.05$). We briefly investigate the sensitivity to the variance parameters controlling the

degree of temporal coherence among sites by considering zero and moderate site differences, with 40 species and the high species turnover pattern for increasing directional change. We only specify two groups of species in our simulation study, but the provided code in Supplement 2 allows for more groups.

RESULTS

Assuming that species composition displayed no directional change, the estimated power was nearly equal to the specified α level for all models. For increasing directional change, as expected, the power of all models increased as the number of years, number of sites, and magnitude of assumed annual univariate linear trend in log odds increased (results not shown). However, for the parameter values we investigated, power was sensitive to the underlying univariate species patterns used to emulate increasing directional change (comparing across rows Fig. 1). The scenario mimicking high species turnover yielded the highest power (last row Fig. 1). Assuming that 50% of species increased over time had less power compared to assuming 50% of species decreased over time, but this effect was minimized as the number of sites increased from 20 to 40 (first to second row Fig. 1). Interestingly, power increased for all models with an increase in the number of species (comparing across graph columns Fig. 1).

Under most scenarios and parameter values we investigated for increasing directional change, CAP had equivalent or higher power compared to PerMANOVA and Mantel (Fig. 1). The finding that PerMANOVA and CAP had higher power compared to the Mantel test is not too surprising given both approaches account for the site-to-site variability in dissimilarity whereas the Mantel test does not. For simulations specifying no differences between sites the power was still higher for the PerMANOVA and CAP compared to Mantel test (results not shown).

Consistent with the results for increasing directional change, the power of all models for detecting decreasing directional change increased with an increase in the number of sites sampled each year and the number of species in the community (Fig. 2). Similarly, for the parameter values we specified, the power of the three models was dependent on the underlying species patterns used to simulate decreasing directional change (comparing across rows Fig. 2). However, in the case of decreasing directional change, the Mantel test had the highest power compared to CAP and PerMANOVA when 50% and 90% of the species were simulated with concave up patterns over time (second and third rows Fig. 2).

Interestingly, the power of CAP and Mantel tests decreased as the proportion of species with concave up temporal trends decreased, but the opposite pattern emerged for PerMANOVA; power was highest if only 10% of the species showed a trend as compared to 50% and 90%. When only 10% of the species (4 or 10 species)

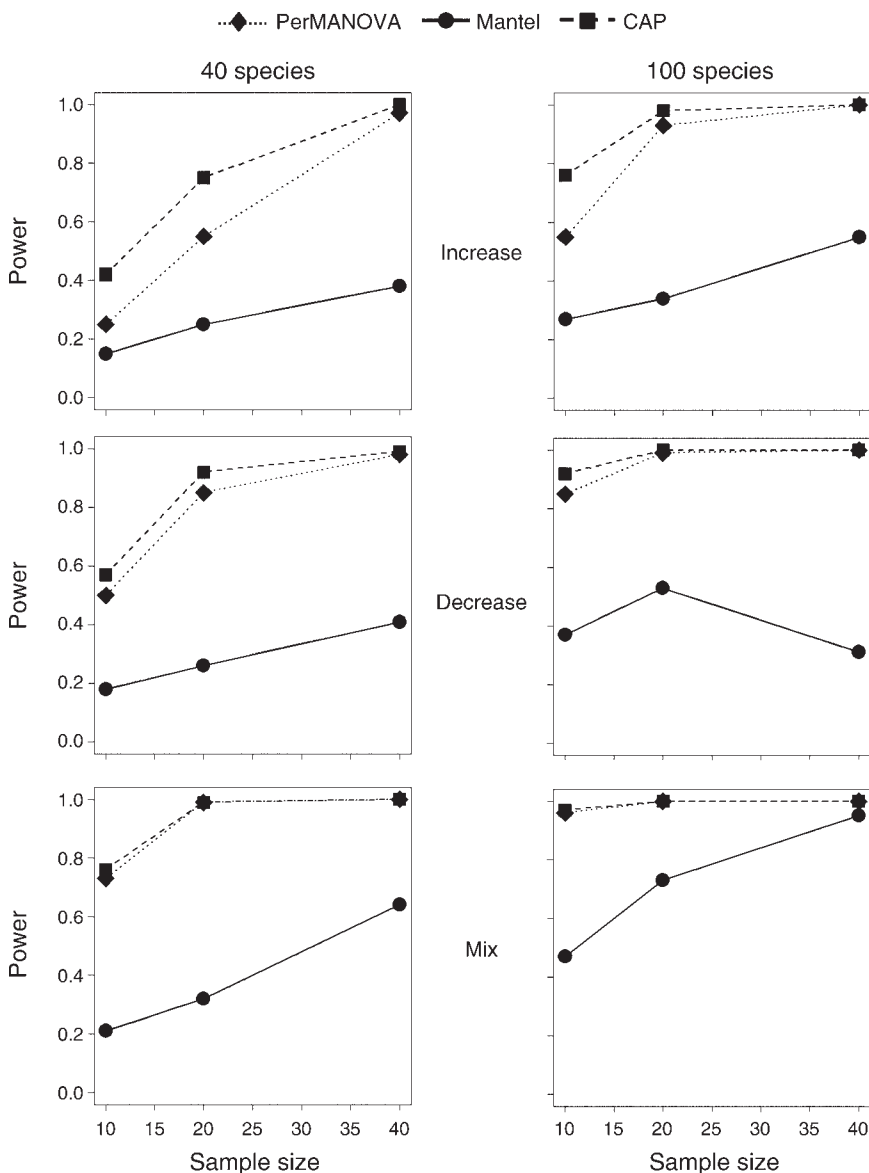


FIG. 1. Estimated power for Mantel, PerMANOVA, and CAP assuming increasing directional change for 40 and 100 species, varying number of sites, three different univariate species patterns over time (first row 50% of species increasing and 50% no trend, second row 50% of species decreasing and 50% no trend, third row 50% increasing and 50% decreasing), 10 years, and $\alpha = 0.05$.

displayed temporal trends CAP was as or slightly more powerful than PerMANOVA, and both were more powerful compared to the Mantel test for a sample size of 40 (first row Fig. 2). The sensitivity to the underlying species pattern became less apparent with 100 species.

DISCUSSION

The number of possible alternative hypotheses in a power analysis to detect trend in a multivariate species assemblage is essentially infinite in terms of the univariate species trends. We found simulating univariate species with quadratic trends generate patterns of decreasing directional change, whereas linear univariate species trends produce patterns of increasing directional

change using Bray-Curtis dissimilarity. This finding narrows the infinite possibilities to a more manageable size. Our investigation into the interplay between number of species, the fraction of the community with significant trend, and different parameter settings was not exhaustive. We did note that for increasing directional change the pattern among sites within time periods is one of convergence; whereas, with decreasing directional change there is a period of divergence followed by convergence among sites within time periods (Appendix A). However, the supplements provide our methods allowing for researchers to investigate their own scenarios of interest. For example, contemporary monitoring programs are emphasizing the

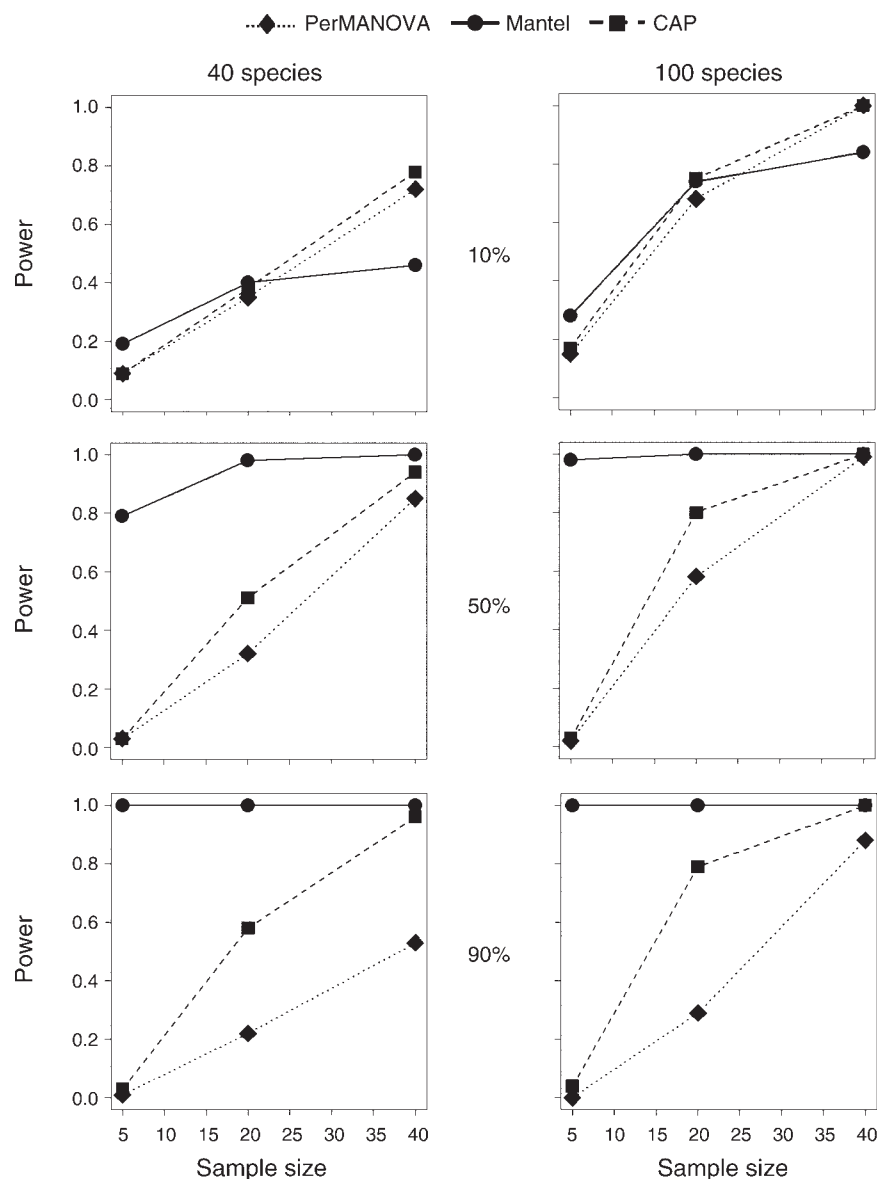


FIG. 2. Estimated power for Mantel, PerMANOVA, and CAP assuming decreasing directional change for 40 and 100 species, varying number of sites, three different univariate species patterns (first row 10% concave up and 90% no change, second row 50% concave up and 50% no change, and third row 90% concave up and 10% no change), 10 years, and $\alpha = 0.05$.

need to articulate or describe climate change scenarios for adaptive management (Thuiller et al. 2005, Millar et al. 2007). Our procedures can be used to investigate whether the sampling effort of a monitoring program is sufficient to detect the articulated changes in species composition under a climate change scenario of interest.

Based on our investigation, CAP was as or more powerful compared to PerMANOVA for increasing directional change. However, for decreasing directional change the Mantel test tended to have higher power, particularly when the majority of species were simulated with concave up temporal trends. One possible explanation is that there was a nonlinear pattern in the pairwise dissimilarity values for the decreasing directional change

scenario (weakly concave down), but we only included a linear term for year in both CAP and PerMANOVA. Power may increase for both CAP and PerMANOVA if a quadratic year term is included. However, we specified the Spearman correlation based on the ranks for the Mantel test which is more powerful for nonlinear monotonic trends. Our results are consistent with Anderson and Robinson (2003) who reported that the power of both a PerMANOVA statistic and the CAP statistic we used were sensitive to the assumed alternative hypothesis. Interestingly, in our simulations the power increased for all models with an increase in the number of species, contrary to the findings by Legendre (2000) for smaller numbers of species (maximum of 50).

For a thorough power analysis, we suggest researchers select a range of reasonable values for the underlying parameters. We briefly investigated the sensitivity of the power estimates for CAP, PerMANOVA, and the Mantel test to the assumed values. We noted a slight decrease in power for an increase in variance parameter values as expected from the knowledge of univariate power analyses (Larsen et al. 2001). Also, increasing the site-level variance components and species pool may decrease the amount of convergence among sites within time periods. Another recommended step is to perform an exploratory data analysis on a subset of the simulated data sets to ensure reasonable conditions are being assumed for a particular application, as we did in Appendix A. If the desired simulated pattern is divergence among sites within years, species by site interactions may need to be incorporated. In other words, species may be increasing or decreasing in their odds of detection/presence over time in only a subset of the sites, which we do not assume in our simulation method.

The benefit of using multivariate tools for detecting trends in composition as compared to a fully multilevel model assuming random species effects (e.g., Dorazio et al. 2010, Gotelli et al. 2010), is that, unless species groups of interest are known a priori, a multi-level model pools the estimates across species which could lead to low power because the overall trend may be zero. For example, the power was highest for all models when 50% of simulated species were increasing and 50% were decreasing over time (increasing average dissimilarity within sites). However, pooling all of the univariate species trend estimates together in a mixed model, in this scenario, would suggest there was no overall trend in the assemblage (Type 2 error). Therefore, we feel for this ecological question of interest (detecting trends in species composition) using multivariate tests based on community dissimilarity measures is appropriate.

ACKNOWLEDGMENTS

We thank S. Mohren and D. Odion for relevant discussions; E. Archer, B. Roper, and S. Miller for access to PIBO data; D. Roberts, T. Rodhouse, S. Hayes, and K. Steinhorst for comments on an earlier version. Two anonymous reviewers significantly helped clarify our terminology for trend in species composition, and we are grateful to both for their assistance. Participation of K. M. Irvine was facilitated through the Pacific Northwest Cooperative Ecosystem Studies Unit cooperative agreement # J8W07060004.

LITERATURE CITED

- Anderson, M. J., and J. Robinson. 2003. Generalized discriminant analysis based on distances. *Australian and New Zealand Journal of Statistics* 45:301–318.
- Anderson, M. J., and T. J. Willis. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84:511–525.
- Angeler, D. G., O. Viedma, and J. M. Moreno. 2009. Statistical performance and information content of time lag analysis and redundancy analysis in time series modeling. *Ecology* 90:3245–3257.
- Collins, S. L., F. Micheli, and L. Hartt. 2000. A method to determine rates and patterns of variability in ecological communities. *Oikos* 91:285–293.
- Dorazio, R. M., M. Kery, J. A. Royle, and M. Plattner. 2010. Models for inference in dynamic metacommunity systems. *Ecology* 2466–2475.
- Eberhardt, L. L., R. R. Knight, and B. M. Blanchard. 1986. Monitoring grizzly bear population trends. *Journal of Wildlife Management* 50:613–618.
- Gotelli, N. J., R. M. Dorazio, A. M. Ellison, and G. D. Grossman. 2010. Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. *Philosophical Transactions of the Royal Society B* 365:3621–3631.
- Kershner, J. L., E. K. Archer, M. Coles-Ritchie, E. R. Cowley, R. C. Henderson, K. Kratz, C. M. Quimby, D. L. Turner, L. C. Ulmer, and M. R. Vinson. 2004. Guide to effective monitoring of aquatic and riparian resources. RMRS-GTR-121. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, USA.
- Larsen, D. P., T. M. Kincaid, S. E. Jacobs, and N. S. Urquhart. 2001. Designs for evaluating local and regional scale trends. *BioScience* 51:1069–1078.
- Legendre, P. 2000. Comparison of permutation methods for the partial correlation and partial mantel tests. *Journal of Statistical Computation and Simulation* 67:37–73.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75:435–450.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology*. Second English edition. Elsevier, Amsterdam, The Netherlands.
- Manley, P. N., W. J. Zielinski, M. D. Schlesinger, and S. R. Mori. 2004. Evaluation of a multiple-species approach to monitoring species at the ecoregional scale. *Ecological Applications* 14:296–310.
- McArdle, B. H., and M. J. Anderson. 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82:290–297.
- Millar, C. I., N. L. Stephenson, and S. L. Stephens. 2007. Climate change and forests of the future: managing in the face of uncertainty. *Ecological Applications* 17:2145–2151.
- Oakley, K. L., L. P. Thomas, and S. G. Fancy. 2003. Guidelines for long-term monitoring protocols. *Wildlife Society Bulletin* 31:1000–1003.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2010. *vegan: community ecology package version 1.17-10*. (<http://cran.r-project.org>)
- Philippi, T. E., P. M. Dixon, and B. E. Taylor. 1998. Detecting trends in species composition. *Ecological Applications* 8:300–308.
- R Development Core Team. 2008. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (www.R-project.org)
- Rosenberg, D. M., and V. H. Resh, editors. 1993. *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York, New York, USA.
- Somerfield, P. J., K. R. Clarke, and F. Olsford. 2002. A comparison of the power of categorical and tests applied to community ecology data from gradient studies. *Journal of Animal Ecology* 71:581–593.
- Thuiller, W., S. Lavorel, M. B. Araujo, M. T. Sykes, and I. C. Prentice. 2005. Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences* 102(23):8245–8250.
- Warton, D. I., and H. M. Hudson. 2004. A MANOVA statistic is just as powerful as distance-based statistics, for multivariate abundances. *Ecology* 85:858–874.

APPENDIX A

Example of simulated communities undergoing increasing and decreasing directional change in species composition over time (*Ecological Archives* E092-159-A1).

APPENDIX B

Table of parameter values used in the simulation study (*Ecological Archives* E092-159-A2).

SUPPLEMENT 1

R code example for fitting the Mantel test, permutational multivariate analysis of variance, and constrained analysis of principal coordinates in the vegan package of R (*Ecological Archives* E092-159-S1).

SUPPLEMENT 2

R code for simulating binary assemblage data that exhibit directional change in species composition (*Ecological Archives* E092-159-S2).