

Ecology, 1998, v.79, n.2, pp.442-447.

ISSN: 0012-9658

DOI: doi: 10.1890/0012-9658(1998)079[0442:FSCTSD]2.0.CO;2

<http://www.esajournals.org/loi/ecol>

[http://www.esajournals.org/doi/pdf/10.1890/0012-9658\(1998\)079%5B0442%3AFSCTSD%5D2.0.CO%3B2](http://www.esajournals.org/doi/pdf/10.1890/0012-9658(1998)079%5B0442%3AFSCTSD%5D2.0.CO%3B2)

© 1998 by the Ecological Society of America

FROM STAR CHARTS TO STONEFLIES: DETECTING RELATIONSHIPS IN CONTINUOUS BIVARIATE DATA

JAMES E. GARVEY,¹ ELIZABETH A. MARSCHALL,² AND RUSSELL A. WRIGHT³

Aquatic Ecology Laboratory, Department of Zoology, Ohio State University, 1314 Kinnear Road, Columbus, Ohio 43212 USA

Abstract. Within many ecological systems, relationships between controlling factors and associated response variables are complex. In many cases, the response should vary little when the controlling factor exerts strong effects. Conversely, when the effect of the controlling factor is weak or absent, the response may vary greatly with effects of other factors. Correlation or regression analyses often may not be appropriate for testing these relationships, because variance of the response changes with values of the controlling factor. We suggest using a technique from the astronomy literature, a two-dimensional Kolmogorov-Smirnov (2DKS) test, to detect relationships in bivariate data with these patterns of variance. This technique successfully identified simulated bivariate data composed of paired independent values as having nonsignificant relationships and simulated bivariate data in which mean and variance of y was constrained at high levels of x as having significant relationships. Using these simulations and examples from aquatic and terrestrial systems, we demonstrate that the 2DKS is a robust test for detecting nonrandom patterns in bivariate distributions that commonly arise in many ecological systems.

Key words: bivariate distributions; correlation; limiting factors; nonparametric test; statistical test; two-dimensional Kolmogorov-Smirnov test; variance.

INTRODUCTION

Within many ecological systems, complex interactions underlie observed patterns. In many cases, functional relationships exist between controlling factors and associated response variables. However, the strength of these relationships often differ greatly over the range of observed values. For example, in many systems, we expect the response to vary little when the controlling factor exerts strong effects. Conversely, when the effect of the controlling factor is weak or absent, the response may vary greatly with the influence of other factors. Though we often attempt to characterize these relationships statistically using correlations and regressions, these unsatisfactory approaches clearly violate the assumption that variance of the response does not change with the magnitude of the controlling factor. These techniques fail to capture an important and interesting characteristic of these relationships: Correlations between the response and controlling factor change over the range of observations.

To illustrate some common features of these types of data sets, we use density data from two co-occurring benthic insects. *Leuctra tenuis* (a stonefly) and *Ameletus ludens* (a mayfly) were sampled in head-

water streams in Ohio (see Dingleline 1996 for details). Though experimental results indicated that *Ameletus* distributions were likely strongly controlled by presence/absence of fish, *Leuctra* densities were not as easily explained by these fish treatments. To answer the post hoc question of whether *Leuctra* might be responding to presence of *Ameletus* rather than to fish, we plotted the densities of *Leuctra* vs. *Ameletus* from each benthic sample (Fig. 1; Dingleline 1996). Clearly, at high *Ameletus* densities, *Leuctra* were rare, whereas at low *Ameletus* densities, *Leuctra* occurred at both high and low densities (Fig. 1). Though *Leuctra* and *Ameletus* densities may have been negatively related when *Ameletus* were abundant, a regression (implying a functional relationship between *Leuctra* and *Ameletus*) over the entire range of data would be inappropriate if other factors drove *Leuctra* densities when *Ameletus* were rare. A traditional measure of correlation (e.g., Pearson product-moment correlation) also would not be appropriate because of its underlying assumption of a linear relationship between variables, which is clearly of little relevance in these data.

Rather than asking how much variance can be explained by these two factors, we suggest asking if the distributions of *Ameletus* and *Leuctra* densities could arise independently. Our initial willingness to interpret the absence of data points in the upper right quadrant of Fig. 1 as indicating that *Ameletus* had a negative effect on *Leuctra* was tempered by the fact that neither species occurred in high numbers very frequently. Even if density distributions of the two species were inde-

Manuscript received 21 November 1996; revised 17 March 1997; accepted 21 March 1997.

¹ Present address: Department of Biology, Queen's University, Kingston, Ontario, Canada K7L 3N6.

² Address correspondence to this author.

³ Present address: Department of Fisheries and Allied Aquacultures, Auburn University, Auburn, Alabama 36849 USA.

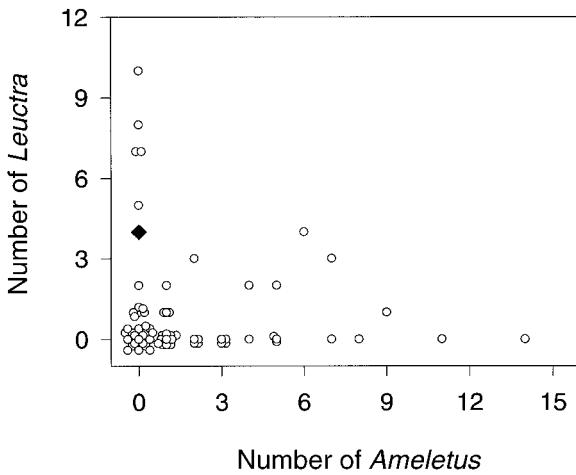


FIG. 1. Relationship between number of *Ameletus* and *Leuctra* per benthic core ($N = 72$ cores) in pools in Ohio headwater streams. About 51% of observations are located at ≤ 1 *Ameletus* and *Leuctra* per benthic core. Symbols representing overlapping values are offset slightly to demonstrate how points are distributed on the graph. The value producing D_{BKS} is denoted by \blacklozenge .

pendent, the probability of observing them jointly at high densities was low, given that high densities of either species rarely occurred. Clearly, a test is required in which we assume that the joint bivariate distribution of densities arose independently, consisting of chance pairings of these data. We then compare the observed joint distribution to this null distribution, determining whether the densities were related.

Recall that the familiar Kolmogorov-Smirnov test compares two univariate cumulative probability distributions by determining if D_{KS} , the maximum difference between the two distributions, is significantly large. We have borrowed a technique from the astronomy literature to extend this approach to bivariate distributions. Because cumulative probability distributions are not well defined beyond one dimension (Press et al. 1992), Fasano and Franceschini (1987) modified an approach by Peacock (1983) to compare two-dimensional (bivariate) distributions. Their solution was to find the maximum difference, D_{BKS} , in integrated probabilities for four quadrants around each point in the plane. If D_{BKS} between an observed and theoretical bivariate distribution (one-sample test) or two observed bivariate distributions (two-sample test) exceeds that expected randomly, we conclude that they differ. To demonstrate the utility of the two-dimensional Kolmogorov-Smirnov (2DKS) test in exploring the types of relationships described above, we provide several examples. First, we use the 2DKS test to detect dependence between two variables in simulated data sets. By then comparing this test to traditional techniques using ecological examples, we demonstrate its potential widespread utility for delineating patterns in ecological data.

METHODS

Two-dimensional Kolmogorov-Smirnov test (2DKS)

To test for independence in a single bivariate data set (one-sample test), we used a 2DKS test (Peacock 1983, Fasano and Franceschini 1987, Press et al. 1992, 1996a, b). For each pair of coordinates ($X_i, Y_i, i = 1, \dots, n$), we counted the points in each of four surrounding quadrants, with X_i, Y_i at the origin,

$$(x \leq X_i, y \geq Y_i) \quad (x > X_i, y \geq Y_i)$$

$$(x \leq X_i, y < Y_i) \quad (x > X_i, y < Y_i)$$

$$(i = 1, \dots, n) \tag{1}$$

and calculated associated proportions. To compare the observed bivariate distribution to one expected if x and y values were distributed independently, for each pair of coordinates we calculated the proportions of observed $x > X_i$, observed $x \leq X_i$, observed $y \geq Y_i$, and observed $y < Y_i$, and multiplied these to get the proportions of (x, y) points that would fall in each quadrant. Within each quadrant, the difference between the observed and expected proportion of points was determined. After all differences were calculated, we determined D_{BKS} as the maximum difference between the observed and expected proportions across all points and quadrants. To determine if D_{BKS} differed from that expected by chance, we re-randomized the original data 5000 times (Manly 1991), performing the 2DKS test each time to get a randomly generated difference, D_{RAN} . Comparing our randomly generated differences to D_{BKS} , we assessed the significance level of the test, $P(D_{RAN} > D_{BKS})$.

Simulated data

In our view, rather than imposing discrete limits to distributions, limiting factors likely restrict their variance and also may affect their mean. To determine whether the 2DKS test detects changes in variance of a dependent variable, y , with respect to a limiting factor, x , we simulated 200 data sets ($N = 250$ points each): 100 in which x and y were distributed independently and 100 in which x restricted variance of y beyond some limit. Values of x and y were generated randomly using a routine from Press et al. (1992), with a different random seed number for each simulation. For both sets of simulations, we generated a random distribution of positive, independent x and y values where small numbers were common but a significant number of large values also were present, using the following generating function from Press et al. (1992):

$$X_i = \left| f + \left(\frac{f}{\sigma_x} \right)^3 \right| \quad Y_i = \left| d + \left(\frac{d}{\sigma_y} \right)^3 \right| \tag{2}$$

where

$$f \approx N(0, \sigma_x^2) \quad d \approx N(0, \sigma_y^2) \tag{3}$$

and $\sigma_x^2 = 100$. For simulations in which values of x

and y were distributed independently, $\sigma_y^2 = 100$ for all values of x . For simulations in which values of y were small, with constrained variance beyond a limit imposed by x ,

$$\sigma_y^2 = \begin{cases} \sigma_{ya}^2 & \text{for } x \leq 10 \\ \sigma_{yb}^2 & \text{for } x > 10. \end{cases} \quad (4)$$

We set $\sigma_{ya}^2 = 100$ and $\sigma_{yb}^2 = 1$.

We then assessed the sensitivity of the 2DKS test to the magnitude of this restriction of variance in y for $x > 10$. Using the same generating function and variance for x as above, we conducted simulations in which standard deviation of y (σ_{yb}) was varied as a proportion of σ_{ya} . We kept $\sigma_{ya} = 10$ and varied σ_{yb} from one (highly constrained) to 10 (i.e., identical distribution of y over the entire range of x) by increments of one. Ten replicates were conducted for each level of variance ($N = 250$ points each; generating a total of 100 data sets).

Empirical examples

We identified potential factors that control variance and abundance of: (1) benthic invertebrates in a stream, (2) fishes in reservoirs, and (3) nectar in flowers. We argue that the 2DKS test should robustly determine the likelihood that these bivariate data derived from two independent, univariate distributions. In contrast, traditional regression techniques should produce poor fits to these kinds of data, where we expect that functional relationships between the variables is limited to some range of values of the limiting factor. (1) We used a 2DKS test to determine whether abundances of *Ameletus* and *Leuctra* were related in stream pools (Fig. 1; Dingleline 1996). We then used linear and nonlinear (assuming a negative exponential fit to the data) least squares regressions to determine if the assumption of a continuous functional relationship between *Ameletus* and *Leuctra* produces satisfactory results (SAS Institute 1985). (2) Age-0 threadfin shad (*Dorosoma petenense*) and gizzard shad (*D. cepedianum*) may compromise growth and survival of other sympatric age-0 fishes such as sunfish (*Lepomis* spp.) and white crappie (*Pomoxis annularis*) through exploitative competition for zooplankton (Guest et al. 1990, DeVries and Stein 1992, Dettmers et al. 1996). To test the hypothesis that *Dorosoma* spp. limit variance and abundance of sunfishes in Ohio reservoirs (<100 ha), we conducted a 2DKS test to determine whether distributions of age-0 *Dorosoma* spp. and sunfish densities collected across four reservoirs and several years (Garvey et al. 1998) could have arisen independently. Again, we applied linear and nonlinear regressions to these data, comparing these results to those of the 2DKS test. (3) In flowers, high nectar volume at a given time may be an indication of generally high productivity; thus we would expect this volume to be positively correlated with production of new nectar following removal of this initial volume. Alternatively, flowers may have a limit to amount of nectar produced over a given period,

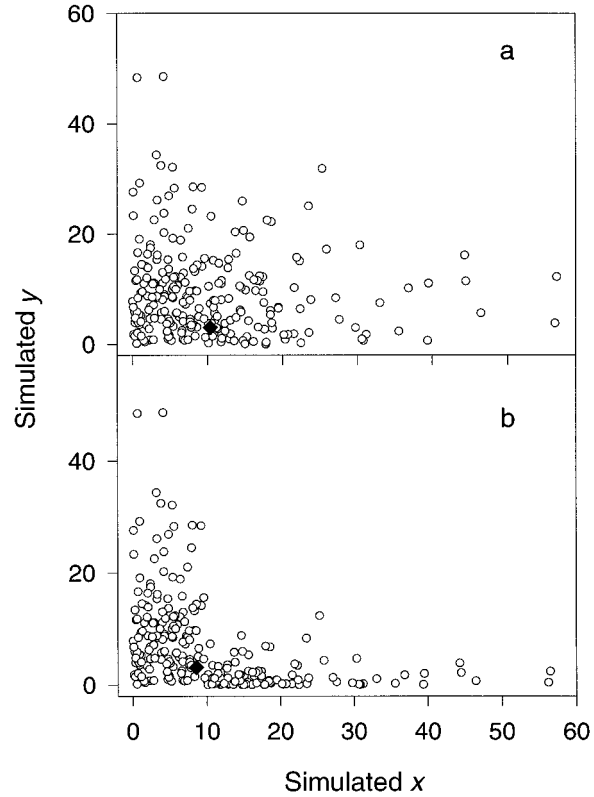


FIG. 2. Example of data from one simulation (total number of simulations = 200; $N = 250$ points) in which x and y values with skewed positive distributions were generated and then tested with a 2DKS test. (a) No pattern present in data; (b) pattern present in data, where, for $x > 10$, variance of randomly generated values of y was constrained. The values producing D_{BKS} are designated by \blacklozenge .

and then we would expect a negative correlation between amount of nectar in a flower at a given time and production of new nectar following removal. Using both the 2DKS test and traditional regression techniques, we tested these expectations in mountain cranberry (*Vaccinium erythrocarpum*), with initial nectar volumes measured by removing all nectar early in the morning (0600–0800) and new nectar production measured in the three hours that followed (on flowers bagged to exclude pollinators; E. A. Marschall and L. Real, unpublished data). All example data sets are available in ASCII format from the second author.

RESULTS

Simulated data

When x and y were distributed independently in simulated data sets, D_{BKS} did not differ significantly from values generated from re-randomized distributions (0.03 ± 0.008 [mean $D_{BKS} \pm 1$ SD]; Fig. 2a). Out of these 100 simulations, only five had nonrandom patterns ($P < 0.05$; 2DKS test), as expected for random data. In simulated data sets in which variance of y was severely restricted at $x > 10$ ($\sigma_{yb}^2 = 0.01 \times \sigma_{ya}^2$), all

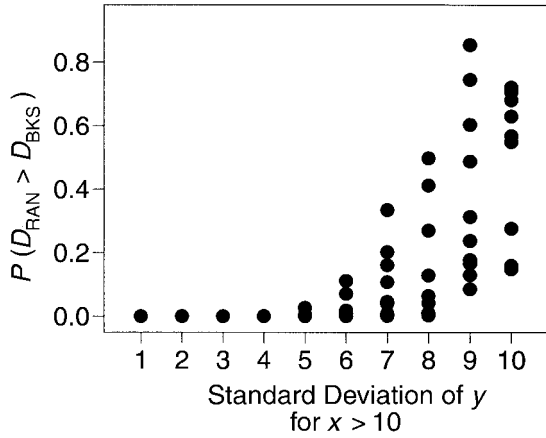


FIG. 3. Sensitivity (i.e., $P(D_{\text{RAN}} > D_{\text{BKS}})$) of the 2DKS test to nonrandom patterns in simulated data, in which standard deviation of y (σ_{y_b}) varied from values of 1 (highly constrained) to 10 (identical distribution of y over the entire range of x) when $x > 10$. Ten replicate data sets were tested at each level of variance ($N = 250$ points each, generating a total of 100 data sets).

100 tests were different from random (mean $D_{\text{BKS}} = 0.13 \pm 0.01$; $P = 0.0002$ for all tests; Fig. 2b). D_{BKS} occurred at x values in the range 8.4–11.4 in these tests (mean $D_{\text{BKS}} = 9.8 \pm 0.6$).

Assessing the sensitivity of the 2DKS test to changes in the magnitude of the restriction of variance in y , we found all patterns were detected as nonrandom below $\sigma_{y_b} = 5$ (2DKS test; $P = 0.0002$ for all tests; Fig. 3). As the restricted standard deviations exceeded 5 (i.e., >25% of unrestricted variance), an increasing proportion of the 2DKS tests failed to detect the simulated nonrandom pattern (Fig. 3).

Examples

In five of six tests, traditional regression produced poor fits for the three example data sets (r^2 ranged from 0.012 to 0.18; Table 1), though all relationships were significant when nonlinear regression was applied (Table 1). Only for the shad–sunfish relationship did nonlinear regression produce a significant and satisfactory fit ($r^2 = 0.75$, Table 1). Though the nonlinear regression results suggested that densities of *Leuctra* and *Ameletus* were weakly related (Table 1), the 2DKS test results revealed that the pattern did not differ from one expected if the univariate distributions for both species

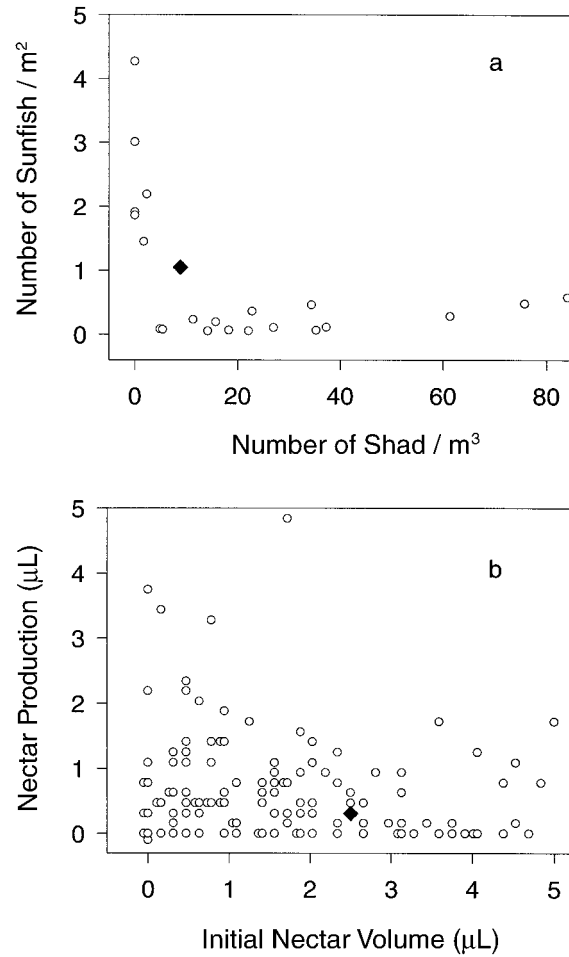


FIG. 4. (a) Relationship between the number of shad/ m^3 captured in offshore zones and number of sunfish/ m^2 seined in inshore areas of four Ohio reservoirs across several years ($N = 22$ lake-years). (b) Relationship between morning nectar volume and nectar produced during three hours following nectar removal in mountain cranberry ($N = 114$ measurements). Symbols representing overlapping values are offset slightly to demonstrate how points are distributed on the graph. The values producing D_{BKS} are denoted by \blacklozenge .

were generated independently ($D_{\text{BKS}} = 0.04$; $P = 0.42$; Fig. 1). In reservoirs, we could reject the null hypothesis of independence of shad and sunfish density ($D_{\text{BKS}} = 0.18$; $P = 0.002$; Fig. 4a); D_{BKS} occurred when $x = 8.8$ shad/ m^3 . In *Vaccinium*, initial morning nectar volume and nectar production through midmorning were

TABLE 1. Results of linear ($Y = b_1X + b_0 + \epsilon$) and nonlinear ($Y = \exp[-b_1X + b_0] + \epsilon$) regressions for example data sets.

Data set	Regression	r^2	df	F	P
<i>Ameletus</i> vs. <i>Leuctra</i> (Fig. 1)	linear	0.012	1, 71	0.85	0.36
	nonlinear	0.028	2, 71	9.74	< 0.05
Shad vs. sunfish (Fig. 4a)	linear	0.18	1, 21	4.6	< 0.05
	nonlinear	0.75	2, 21	53.3	< 0.05
Nectar (Fig. 4b)	linear	0.039	1, 113	4.6	< 0.05
	nonlinear	0.043	2, 113	93.9	< 0.05

related ($D_{\text{BKS}} = 0.07$; $P = 0.02$; Fig. 4b), with D_{BKS} occurring when initial volumes were 2.5 μL .

DISCUSSION

Traditional tools used to assess patterns in bivariate data (e.g., regressions and traditional correlation analyses) are only appropriate for particular kinds of data. For example, traditional correlation analyses are built on the assumption that an underlying linear correlation exists between two variables. And regression analyses are appropriate for data in which there is a functional relationship between x and y over the entire range of observed values of these variables. Indeed, though regression results for our example data were often significant, generally poor regression model fits to the data, plus the lack of a functional relationship at some range of values of the limiting factors, support our view that such approaches are inappropriate for these kinds of data. Hence, our intention is not to suggest an alternative approach to traditional techniques, but rather to offer a tool (i.e., the 2DKS test) that will allow us to analyze these different kinds of data in which (1) we do not expect a linear correlation between two variables, and (2) we expect that any functional relationship between the variables breaks down over some range of values of one of the variables, and thus we expect high variance in the other variable across this range.

The 2DKS test detected nonrandom patterns in simulated, bivariate data, in which a fixed factor constrained variance beyond a threshold point. Of course, our ability to detect these nonrandom patterns depended on the relative constraint of variance of y beyond the threshold. In simulations in which variance of y at values of $x > 10$ was 0.01 of unconstrained variance, D_{BKS} consistently occurred at values of x close to that threshold point. Hence, if changes in variance occur at relatively discrete points, the 2DKS test will detect these fixed threshold points. When biotic or abiotic factors control responses only beyond some critical level of these factors, this test may be valuable to detect and estimate the threshold level. More commonly, though, variance of a dependent variable may not be constrained at some fixed threshold value of the independent variable. Rather, the bivariate distributions will vary in complex ways, generating shapes that deviate substantially from normal bivariate distributions. Because the 2DKS test is distribution-free (Peacock 1983; Fasano and Franceschini 1987), it should be robust to nonrandom patterns in bivariate distributions common to ecological systems.

Extending the utility of the 2DKS test beyond simulated data, we determined whether bivariate distributions from ecological systems were produced randomly. As with our simulated data, low densities of *Ameletus* and *Leuctra* were common whereas high densities were relatively rare. Though *Leuctra* densities appeared to decline when *Ameletus* was present, the 2DKS test demonstrated that this pattern could easily

have arisen from their independent univariate distributions. Thus, the author cannot conclude that interspecific interactions between these benthic insects caused observed densities of *Leuctra*. Other factors, perhaps chemical cues from fishes in upstream pools, contributed to observed distributional patterns (Dingledine 1996). In our view, the 2DKS test allowed us to objectively determine that this bivariate distribution was not different from random and, subsequently, to pursue other potential ecological mechanisms.

Unlike invertebrates in streams, field distributions of reservoir fishes and nectar produced by flowers likely were not generated by chance. In Ohio reservoirs, experimental and field evidence suggest that densities >10 age-0 shad/ m^3 may limit growth and survival of sympatric age-0 sunfishes (Stein et al. 1995, Dettmers and Stein 1996, Dettmers et al. 1996, Garvey et al. 1998). Though our nonlinear regression explained significant variance in the shad–sunfish relationship, we believe that this model may be misleading because it implies a continuous relationship, whereas in reality shad and sunfish abundances may be unrelated for densities <10 age-0 shad/ m^3 . In support of this view, the 2DKS test detected D_{BKS} when $x = 8.8$ shad/ m^3 , suggesting that shad may only constrain the mean and variance of sunfish densities beyond this abundance. In cases such as these, the 2DKS test may be appropriate for identifying threshold levels in ecological data. In contrast, the significant bivariate distribution produced by initial nectar volume and nectar production probably was not generated when nectar production was constrained beyond some discrete, fixed abundance. This distribution probably arose as a result of some limit to total nectar production per flower in a given morning and among-flower variance in the magnitude of this production. Finding statistical significance in this relationship between these variables gives us cause to further pursue the underlying processes.

Our examples demonstrate the efficacy of the 2DKS test for exploring patterns within complex ecological data sets. This procedure is computationally simple, requiring relatively short times to perform the test and associated 5000 re-randomizations (e.g., <10 min for a data set with 250 observations). As noted by Thomson et al. (1996), who explored associations between variables in complex spatial data, tests for identifying patterns in many bivariate data are largely descriptive. They suggested, without providing examples, that the 2DKS test should provide a robust quantitative test for bivariate distributions. Indeed, this test worked well with bivariate data common to ecological systems. We caution that this test probably will not perform well with data distributed in a fashion that produces multiple, equivalent points of maximum difference. This may occur when points occur as highly clumped patches uniformly distributed in the x – y plane.

Though we only provided tests of independence for bivariate distributions from single samples, the 2DKS

test can be used to compare bivariate distributions from two samples as well. This two-sample test can be quite useful when comparing distributions generated by two experimental treatments or testing theoretical or simulated distributions with actual observations. Because ecological relationships often extend beyond two-dimensional distributions, this test can be modified to compare three-dimensional distributions, again producing statistically robust results (Fasano and Franceschini 1987).

ACKNOWLEDGMENTS

Support for this research was provided to E. A. Marschall by National Science Foundation (NSF) grant DEB 9410327, to J. E. Garvey and R. A. Wright by NSF grant DEB 9407859 and Federal Aid in Sport Fish Restoration Project F-69-P (administered jointly by the U.S. Fish and Wildlife Service and the Ohio Division of Wildlife, to R. A. Stein, The Ohio State University), and to J. E. Garvey by a Presidential Fellowship from The Ohio State University. We thank A. R. Solow and two anonymous reviewers for helpful comments that improved the manuscript.

LITERATURE CITED

- Dettmers, J. M., D. R. DeVries, and R. A. Stein. 1996. Quantifying responses to hybrid striped bass predation across multiple trophic levels: implications for reservoir biomanipulation. *Transactions of the American Fisheries Society* **125**:491–504.
- Dettmers, J. M. and R. A. Stein. 1996. Quantifying linkages among gizzard shad, zooplankton, and phytoplankton in reservoirs. *Transactions of the American Fisheries Society* **125**:27–41.
- DeVries, D. R. and R. A. Stein. 1992. Complex interactions between fish and zooplankton: quantifying the role of an open-water planktivore. *Canadian Journal of Fisheries and Aquatic Sciences* **49**:1212–1227.
- Dingledine, N. A. 1996. The effects of fish on invertebrates: the importance of spatial location in streams. Thesis, Ohio State University, Columbus, Ohio, USA.
- Fasano, G., and A. Franceschini. 1987. A multidimensional version of the Kolmogorov-Smirnov test. *Monthly Notices of the Royal Astronomical Society* **225**:155–170.
- Garvey, J. E., N. A. Dingledine, N. S. Donovan, and R. A. Stein. 1998. Exploring spatial and temporal variation within reservoir food webs: predictions for fish assemblages. *Ecological Applications* **8**:104–120.
- Guest, W. C., R. W. Drenner, S. T. Threlkeld, F. D. Martin, and J. D. Smith. 1990. Effects of gizzard shad and threadfin shad on zooplankton and young-of-year white crappie production. *Transactions of the American Fisheries Society* **119**:529–536.
- Manly, B. F. J. 1991. *Randomization and Monte Carlo methods in biology*. Chapman and Hall, New York, New York, USA.
- Peacock, J. A. 1983. Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society* **202**:615–627.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical recipes in C: the art of scientific computing*. Cambridge University Press, New York, New York, USA.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1996a. *Numerical recipes in FORTRAN 77: the art of scientific computing*. Volume 1. Cambridge University Press, New York, New York, USA.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1996b. *Numerical recipes in FORTRAN 90: the art of scientific computing*. Volume 2. Cambridge University Press, New York, New York, USA.
- SAS Institute. 1985. *SAS users guide: statistics, version 5 edition*. Statistical Analysis System Institute, Cary, North Carolina, USA.
- Stein, R. A., D. R. DeVries, and J. M. Dettmers. 1995. Food web regulation by a planktivore: exploring the generality of the trophic cascade hypothesis. *Canadian Journal of Fisheries and Aquatic Sciences* **52**:2518–2526.
- Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro, and P. Legendre. 1996. Untangling multiple factors in spatial distributions: lilies, gophers, and rocks. *Ecology* **77**:1698–1715.