Letters to the Editor

Identifying determinants of bird diversity

13 April 2009

To the Editor:

Studying five sites in the northwestern United States, Verschuyl et al. (2008) draw conclusions about the influences of available energy and forest structural complexity on bird species richness. Reanalyzing data from one of their figures, kindly provided by the authors, I found little or no support for several of their inferences. A combination of calculation errors and a failure to include site as a factor in their statistical models appear to be responsible for these discrepancies.

Conclusion 1

The authors' conclusion in their abstract that "richness had a hump-shaped relationship with available energy across the northwestern United States" is based on curves fit to plots of species richness against gross primary productivity (GPP) and normalized difference vegetation index (NDVI) in their Fig. 3. My Fig. 1 reproduces these scatterplots and shows, with dashed lines, the curves fit by the authors: a breakpoint regression for the GPP plot, and a cubic regression for the NDVI plot.

The authors' statistical modeling failed to include site as a predictive factor. This would be warranted only if we knew that the five sites were identical with respect to richness-influencing factors other than the one(s) included in the model: an assumption that is obviously untenable, given the multitude of biotic and abiotic factors thought to influence bird species richness (MacArthur and MacArthur 1961, Hawkins and Porter 2003). One manifestation of this problem is the conspicuous structure in plots of residuals from the authors' fitted models (my Fig. 2): residuals are much more similar within sites than across sites, violating the independence assumption that is a cornerstone of inference based on these models.

Table 1 summarizes a variety of models fit to these data, including models that incorporate site as a factor. By the authors' criterion of minimizing Akaike's information criterion (AIC), the "best" sets of predictors are (1) the main effects of site and GPP, and (2) site, NDVI, and the interaction of site and NDVI. The former model implies that the slope of richness vs. GPP is the same at all five sites, and the latter model implies that the slope of richness vs. NDVI varies among sites (see my Fig. 1).

When site is included as a predictor, there is no evidence for a hump-shaped relationship between species richness and GPP, contrary to the breakpoint model fit by Verschuyl et al. (2008). Rather, there is strong evidence that the sites differ in mean species richness, and that, within sites, species richness tends to increase with GPP.

The model containing site, NDVI, and their interaction suggests that slopes tend to decrease from left to right in my Fig. 1B, consistent with the inner portion of the cubic regression fit by the authors, which lacks site as a predictor. However, a closer look shows that only one site, Springfield, has a slope of richness vs. NDVI that is different from those in other sites (Yellowstone, Gold Fork, and Cle Elum; see Fig. 3).

Some of the variability in mean species richness among sites could well be "caused" by differences in the mean energy availability among sites, but the site-level "effect" of energy availability cannot be distinguished from the within-site "effects" of energy availability using this design. The large difference in mean species richness between Springfield and the Coast Range—sites having similar values of the energy predictors—provides evidence of site effects that are unrelated to energy availability (see my Fig. 3). Consequently, the most useful statistical approach to these data is to summarize the within-site trends, as the authors do in reaching their next conclusion.

Conclusion 2

The authors state that "the landscape-scale relationships between energy and richness were positive or hump-shaped in energy-limited locations and were flat or negative in energy-rich locations." This conclusion is based on examination of site-specific relationships between richness and GPP and NDVI.

Fig. 3 shows confidence intervals for the slopes of linear regressions of richness vs. GPP and NDVI, done separately for the five sites. For GPP (Fig. 3A), the only statistically significant slopes are for Gold Fork and Springfield. The most energy-rich site, the Coast Range, has a slope very close to zero. The confidence intervals for the slopes of richness vs. NDVI follow a pattern closer to that claimed by the authors (Fig. 3B), although, again, the most energy-rich site (the Coast Range) has a slope that is statistically indistinguishable from zero.

The authors' evidence for conclusion 2 is summarized in their Table 8, which is fraught with errors, or at least with disparities from my results. My Table 2 summarizes
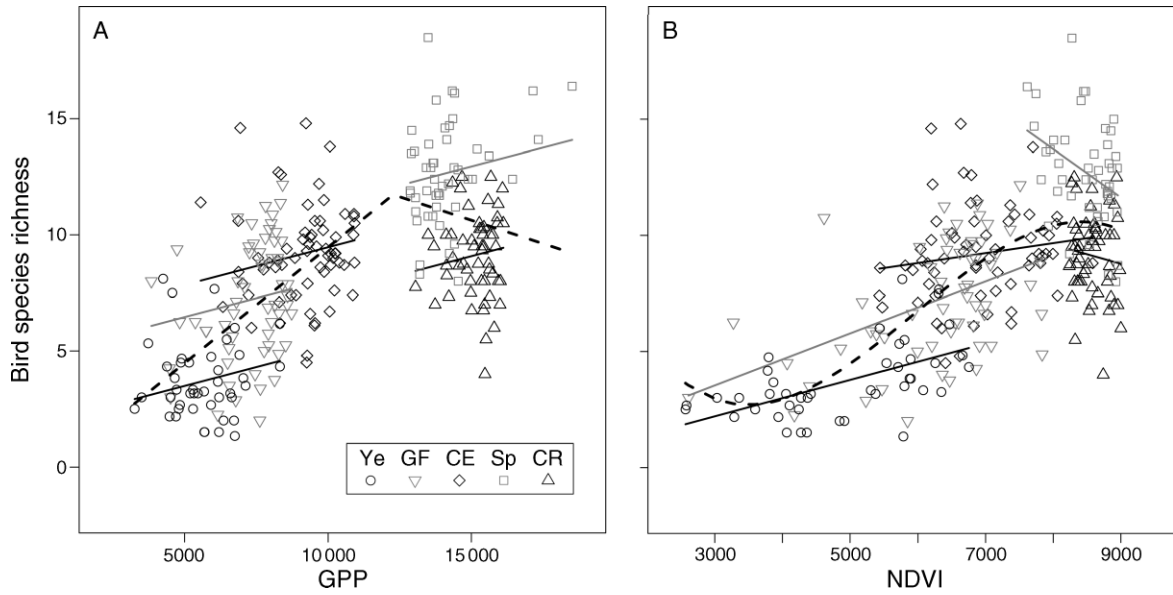
FIG. 1. Plots of bird species richness vs. (A) gross primary productivity (GPP), and (B) normalized difference vegetation index (NDVI), with symbols indicating the different sites (Ye, Yellowstone; GF, Gold Fork; CE, Cle Elum; Sp, Springfield; CR, Coast Range). In each plot, the dashed black lines are the fits preferred by Verschuyl et al. (2008), and the shorter segments are site-specific fits preferred by Murtaugh (see Table 1).

these disparities; our results agree in only three of the 10 cells in the table. In addition, Verschuyl et al. (2008) claimed that cubic and quadratic regressions of richness vs. GPP were warranted for Cle Elum and Gold Fork, respectively, but I found no evidence for nonlinearity in either of these relationships, based on extra-sum-of-squares tests or the AIC.

Conclusion 4

The authors conclude that "the slope of the relationship between forest structural complexity and richness was steepest in energy-limited locations." This observation is based on regressions of bird species richness against an index of variation in tree size (SCI) for three
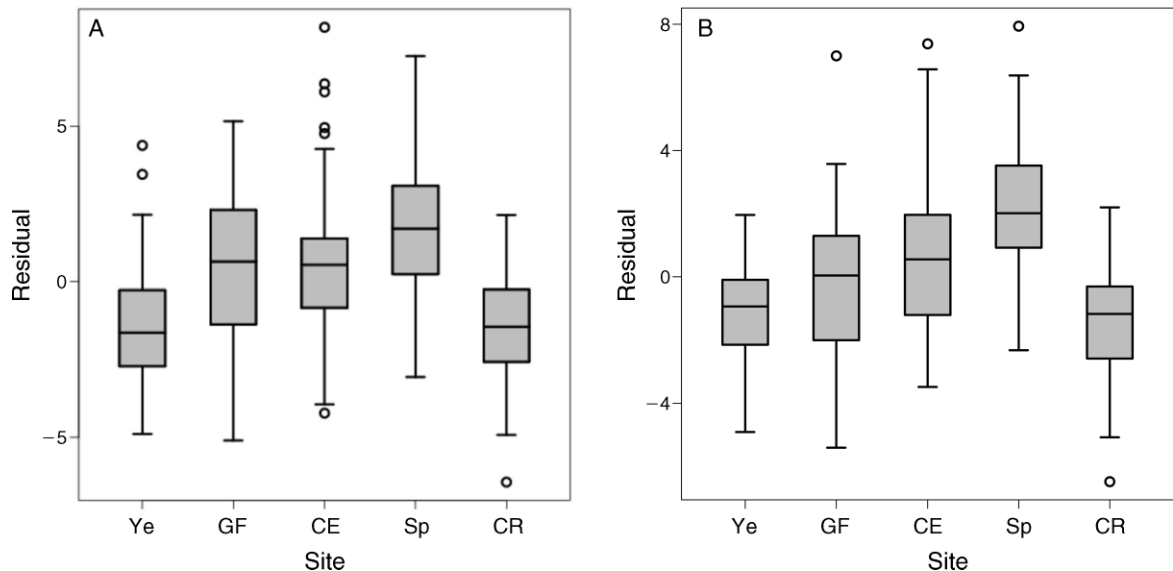

FIG. 2. Box plots of residuals from the species-richness regressions preferred by Verschuyl et al. (2008): (A) a breakpoint model of GPP and (B) a quadratic model of NDVI. Site abbreviations are as in Fig. 1. Each box shows the median and the upper and lower quartiles of the response. The whiskers mark the most extreme observations whose distances to the top or bottom of the box are within 1.5 times the interquartile range; more extreme points are plotted individually.

TABLE 1. Regression models of species richness vs. gross primary productivity (GPP) and normalized difference vegetation index (NDVI), with values of Akaike's information criterion (AIC) and the coefficient of determination.

Model	GPP		NDVI	
	AIC	R^2	AIC	R^2
Site $+ x$	1168.4†	0.64	1157.1	0.65
Site $\times x$	1170.0	0.65	1145.6†	0.68
Site $+ x^2$	1169.7	0.64	1152.5	0.66
Site $\times x^2$	1175.2	0.65	1153.4	0.68
Breakpoint	1272.3‡	0.46	1263.6	0.48
x^2	1278.9	0.44	1269.9	0.46
x^3	1280.6	0.44	1258.4‡	0.49

Notes: Models with only main effects have "+" signs; "×" indicates interactions; and exponents indicate the highest order of the GPP or NDVI terms. For example, the model in the fourth line would have an intercept, four indicators for site, linear and squared terms in x, and all pairwise products of the site and x terms.

† Preferred by P. Murtaugh.
‡ Preferred by Verschuyl et al. (2008).

sites, in their Fig. 6. The slopes for Cle Elum and Yellowstone (sites that are low to intermediate on the authors' energy-availability scale) are positive and statistically indistinguishable, while the slope for Springfield, with ostensibly more available energy, is negative. Using three numbers to draw conclusions about how the relationship between two factors (species richness and SCI) varies among levels of a third (energy availability) is unwarranted, in my opinion.

General comments

Implicit in the authors' analyses and conclusions are the assumptions that (1) the five sites in the paper are representative of some larger population of sites, and (2) the differences in bird species diversity among sites are largely caused by differences in energy availability and forest structural complexity. The paper focuses on the nature of the relationships between species richness and these two "drivers" and argues that these patterns apply to gradients of energy and structural complexity in some larger population of sites.

With respect to the first assumption, it is well known that inference based on nonrandom samples may lead to grossly inaccurate summaries of patterns in the population from which the samples were taken (e.g., see Peterson et al. 1999). This study obviously could not be based on a random sample of landscapes in the northwestern United States, given the intensity of sampling that would be required to match the availability of data for the five studied locations. Consequently, inferences must be based on within-site analyses, and generalization to other sites is wholly speculative.

With respect to the second assumption, it is an axiom of data analysis that correlation does not imply causation, and that the only irrefutable evidence of causation comes from randomized experiments (e.g., see Ramsey

and Schafer 2002:5–7). Experiments are obviously impossible in this and many other ecological contexts, but that argues for special care in interpreting the available observational evidence. In their *Discussion*, Verschuyl et al. (2008:1167) conclude that "Our results suggest that forest structure and available energy drive species diversity with varying strengths throughout the northwestern United States," and they follow with a section on the implications of their findings for biodiversity management. Of the multitude of biotic, abiotic, and historical factors known to influence species diversity, why are the predictors focused on here (GPP and NDVI) accorded the special status of key "drivers" of biodiversity?
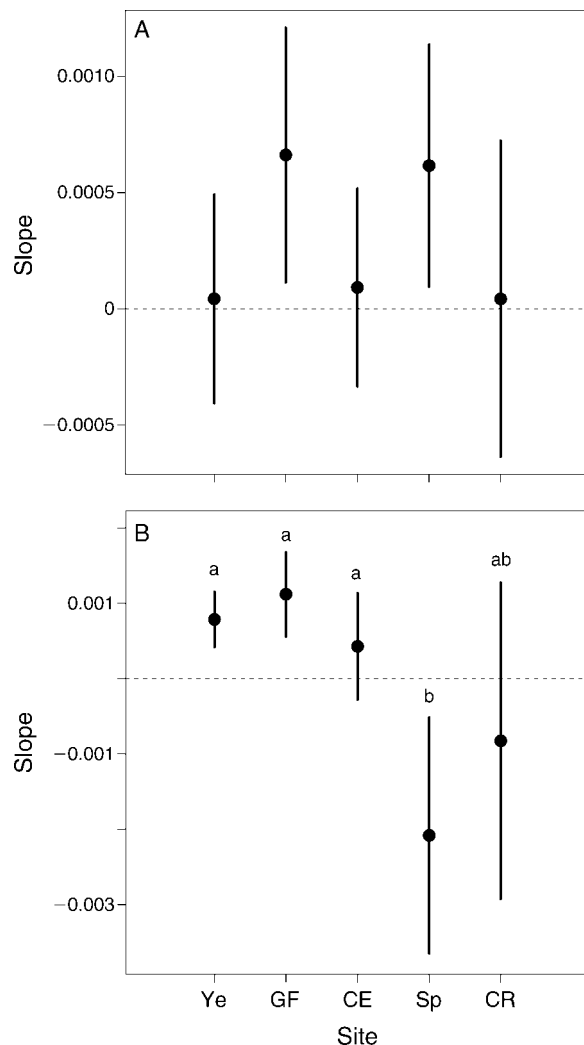

FIG. 3. Estimates of, and 95% confidence intervals for, the slopes of site-specific regressions of species richness vs. (A) GPP and (B) NDVI. In (A), none of the slopes is statistically distinguishable from the others; in (B), intervals labeled with the same letter, a or b, are not statistically different at the $P < 0.05$ level.

TABLE 2. Direction (+, −) and statistical significance of the slopes of linear regressions of species richness vs. GPP and NDVI, as calculated by Verschuyl et al. (2008), indicated by a "v," and by Murtaugh ("pm").

Predictor	Coast Range	Springfield	Cle Elum	Gold Fork	Yellowstone
GPP (v)	ns	ns	+***	+***	+***
GPP (pm)	ns	+*	ns	+*	ns
NDVI (v)	−*	−**	+**	+***	+***
NDVI (pm)	ns	−*	ns	+***	+***

*$P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, not significant.

A potential pitfall of inferring cause and effect from an observational study is that a confounding factor, i.e., a third factor associated with the supposed causative factor and with the response, could just as easily explain the observed association. In extreme cases, confounders can even reverse the apparent direction of the association between two variables, a phenomenon known as Simpson's paradox (e.g., see Wagner 1982). In an example similar to my Fig. 1, Link and Sauer (1998) showed that apparent temporal changes in the abundance of house sparrows along a Breeding Bird Survey route were wholly accounted for by changes in observers over that period.

In my opinion, the relationships of bird species richness to GPP and NDVI are inadequately summarized by Verschuyl et al. (2008); the claimed trends are extrapolated and generalized well beyond the limitations imposed by the study design; and their conclusion that GPP and NDVI "drive" species richness in the northwestern United States is unwarranted.

Literature cited

Hawkins, B. A., and E. E. Porter. 2003. Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America. Global Ecology and Biogeography 12:475–481.

Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological Applications 8:258–268.

MacArthur, R. H., and J. W. MacArthur. 1961. On bird species diversity. Ecology 42:594–598.

Peterson, S. A., N. S. Urquhart, and E. B. Welch. 1999. Sample representativeness: a must for reliable lake condition estimates. Environmental Science and Technology 33:1559–1565.

Ramsey, F., and D. Schafer. 2002. The statistical sleuth: a course in methods of data analysis. Second edition. Duxbury, Pacific Grove, California, USA.

Verschuyl, J. P., A. J. Hansen, D. B. McWethy, R. Sallabanks, and R. L. Hutto. 2008. Is the effect of forest structure on bird diversity modified by forest productivity? Ecological Applications 18:1155–1170.

Wagner, C. H. 1982. Simpson's paradox in real life. American Statistician 36:46–48.

PAUL A. MURTAUGH

Department of Statistics
Oregon State University
Corvallis, Oregon 97331 USA
E-mail: murtaugh@science.oregonstate.edu

Response:

8 April 2009

Murtaugh's criticism of Verschuyl et al. (2008) focuses on whether site should be a covariate in the analysis of the relationship between bird species richness (BSR) and available energy. His perspective largely ignores the state of knowledge on the science question posed and misinterprets the inference that may be drawn from correlational relationships. His critique is also weakened by use of an incomplete data set. He was sent a subset of the data to fulfill his request for data for a class project.

The relationship between species richness and available energy has been the subject of considerable interest and controversy on the nature of the relationships and the potential causal factors (Hansen and Rotella 1999, Irwin 1999, Waide et al. 1999, Hurlbert 2004, Monkkonen et al. 2006, Reed et al. 2006, Bai et al. 2007, Qian et al. 2007, and many others). Most studies have been done at national to continental scales using species occurrence records or national taxon surveys (Mittelbach et al. 2001).

In the Pacific Northwest United States (PNW), many studies have documented a relationship between species richness and vegetation structure. This relationship has strongly guided forest management strategies for biodiversity. Our goal was to examine the amount of variation in BSR explained by energy compared to vegetation structure and to draw implications for management. Ours was the most comprehensive field study to date of the interaction of between energy and structure at regional scales. We sampled 600 points in 134 stands across three landscapes and drew on previous data for two additional landscapes. The full data set include 270 stands within five landscapes arrayed along the energy gradient from the Pacific Coast to the east side of the Rockies.

We asked how much variation in BSR is explained by measures of energy (1) across the region and (2) within each landscape? (3) How much variation is explained by vegetation structure at these two scales? (4) To what extent do energy and structure interact in their relation-

ship with BSR? The initial two analyses are separated by scale (regional and landscape). We offered methods that accounted for the different scales to tackle complicated questions of interest. In his critique of this work, Murtaugh has combined scales and offers a far less informative set of methods to tackle the questions of interest. We used model selection approaches to evaluate several linear and curvilinear models. We found at the regional scale that the best model for energy and richness was unimodal ($r^2 = 0.44$ [annual] and 0.49 [breeding-season]). This result was supported by landscape scale analyses showing positive slopes in lower energy landscapes and flat or negative slopes in higher energy landscapes (Verschuyl et al. 2008: Table 8).

Murtaugh's main criticism of the work is that site was not a covariate in the regional scale analysis. The purpose of our regional analysis was to examine large-scale effects, and the study landscapes used to represent the regional gradient were intentionally selected to provide a broad range of energy input values. Using sites as main effects would have prevented us from modeling the energy effects across the full spectrum of energy values. Because our five landscapes overlapped relatively little in energy level, it is impossible to determine with this extensive data set what proportion of the site effect is due to available energy. Hence, we interpreted the correlation between BSR and energy at the regional level, and supporting relationships at the landscape level, to be evidence supporting a unimodal species energy relationship. Including site as a covariate subsumes much of the variation explained by energy in the univariate model. We feel that reporting this site effect is important. However, energy and vegetation structure are likely candidates to account for much of this site effect. We focused on evaluating each of these two factors directly. As pointed out in the paper, these results lead to the conclusion that more expensive approaches, such as experiments, are justified on energy and structure based on these correlational results.

Murtaugh is correct that his Table 2 and the original Table 8 from our work have many disparities. These disparities, however, stem from Murtaugh not having the full data set to test the landscape scale analyses completed by Verschuyl et al. (2008). In our landscape scale analyses, breeding season normalized difference vegetation index (NDVI) was sampled at four different times during the breeding season. Only the average NDVI value used for the regional analysis was provided in the dataset given to Murtaugh. For the landscape scale analyses, the date that NDVI was collected for each site varied in order to fully expand (and make more meaningful) the gradient of breeding season energy at each site. Therefore, in the mesic Coast Range and Springfield sites where the peak of vegetation growth occurs in early spring, a measure of NDVI from early May was used to assess the relationship with BSR. In

Yellowstone, peak growth does not occur until early summer and therefore we used a measure of NDVI from late June. In the regional analysis these multiple measures were combined into a breeding season average NDVI (the covariate given to Murtaugh for his class project). These differences highlight both the flexibility of methods we used for the two scales analyses, and the imprudence of the critique by Murtaugh.

We point out at length that this is a correlational study and that the resulting relationship between energy and richness is consistent with a causal relationship, but insufficient to demonstrate a causal relationship. The care to not overstep the results in implying inference was clearly stated in the original text in several locations, most notably in the *Discussion* on page 1164: "Although several of the results suggest a downturn in species richness at the highest levels of available energy, *factors other than the availability of energy itself may play a role in reducing species richness. Other potential explanations include forest structure and canopy closure differences inherent to specific landscapes, the distance from the coast, or the physiological limitations of vegetation that may result in an inability to utilize additional energy.*"

The value of the work is that it is the first to demonstrate a correlational relationship between BSR and energy across the PNW and to compare the strength of this with that based on the well tested vegetation structure relationship. We conclude that more work is necessary to establish a causal relationship and point out implications for management.

In concluding his review, Murtaugh suggests that predictors gross primary productivity (GPP) and NDVI were incorrectly singled out and "… accorded the special status of key 'drivers' of biodiversity." A more cautious review of our work and a more complete understanding of the literature would show that GPP and NDVI were not singled out as the key drivers of biodiversity, but were simply used to represent the gradient of available energy across the region as is commonly done in species energy literature (Hansen and Rotella 1999, Irwin 1999, Waide et al 1999, Mittlebach et al. 2001). We took care in selecting both a breeding season and annual measure of available energy (which is somewhat rare in species energy literature), and did not use potentially confounding variables such as precipitation or seasonal measures of temperature in the regional analysis of the relationship between richness and energy.

Murtaugh's review of the work completed by Verschuyl et al. reflects a potential lack of familiarity with the literature on species-energy theory, and a misguided attempt to overstate the cautious conclusions outlined in Verschuyl et al. (2008). In addition, while conducting his analyses with only a subset of the full data, Murtaugh only offered a simplified and uninformative alternative to the analyses we conduct. Thus, we firmly stand by the conclusions of Verschuyl et al. (2008).

Literature cited

Bai, Y. F., J. G. Wu, Q. M. Pan, J. H. Huang, Q. B. Wang, F. S. Li, A. Buyantuyev, and X. G. Han. 2007. Positive linear relationship between productivity and diversity: evidence from the Eurasian Steppe. Journal of Applied Ecology 44: 1023–1034.

Hansen, A. J., and J. J. Rotella. 1999. Abiotic factors. Pages 161–209 *in* M. L. Hunter, editor. Maintaining biodiversity in forest ecosystems. Cambridge University Press, Cambridge, UK.

Hurlbert, A. H. 2004. Species–energy relationships and habitat complexity in bird communities. Ecology Letters 7:714–720.

Irwin, L. L. 1999. Abiotic influences on bird-habitat relationships. Pages 209–218 *in* J. M. Marzluff and R. Salabanks, editors. Avian conservation: research and management. Island Press, Washington, D.C., USA.

Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381–2396.

Monkkonen, M., J. T. Forsman, and F. Bokima. 2006. Energy availability, abundance, energy-use and species richness in forest bird communities: a test of the species energy theory. Global Ecology and Biogeography 15:290–302.

Qian, H., X. Wang, S. Wang, and Y. Li. 2007. Environmental determinants of amphibian and reptile species richness in China. Ecography 30:471–482.

Reed, A. W., G. A. Kaufman, and D. W. Kaufman. 2006. Species richness–productivity relationship for small mammals along a desert–grassland continuum: differential responses of functional groups. Journal of Mammalogy 84(4): 777–783.

Verschuyl, J. P., A. J. Hansen, D. B. McWethy, R. Sallabanks, and R. L. Hutto. 2008. Is the effect of forest structure on bird diversity modified by forest productivity? Ecological Applications 18:1155–1170.

Waide, R. B., M. R. Willig, S. F. Steiner, G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday, and R. Parmenter. 1999. The relationship between productivity and species richness. Annual Review of Ecology and Systematics 30:257–300.

J. P. Verschuyl,[1] A. J. Hansen, and D. B. McWethy

Hamer Environmental
1510 S. 3rd Street
P.O. Box 2561
Mount Vernon, Washington 98273 SA

[1] E-mail: jverschuyl@yahoo.com