The exclusion of rare species from community-level analyses

Dianna M. Queheillalt, James W. Cain III, Denise E. Taylor, Michael L. Morrison, Stacia L. Hoover, Natasha Tuatoo-Bartley, Lourdes Rugge, Kirsten Christopherson, Miriam D. Hulst, Michael R. Harris, and Heather L. Keough

Abstract An accurate assessment of animal abundance is essential to understand the effects of natural variables or anthropogenic activities on wildlife communities. Many studies attempting to assess the effects of such impacts on wildlife communities use only a portion of species present to render management recommendations for an entire community. Abundance is often not calculated for all species detected, and species are excluded from community-level analyses due to their rarity. To determine how frequently authors excluded species from their analyses, we reviewed 190 articles from 11 wildlife-related journals published from 1983–2001. Of the articles we reviewed, 31% did not include all species in their analyses due to small sample sizes. Of the articles excluding species from calculations, abundances were not calculated for an average of 50% of the species detected due to small sample sizes. These studies made management recommendations for entire communities or areas based on <50% of the species present. We suggest alternatives to excluding species from community-level analyses. By reducing the number of species excluded from community-level analyses due to small sample sizes, researchers can make more accurate assessments of the effects of natural and anthropogenic factors on wildlife communities.

Key words abundance, community, density, diversity, rare species

Research in conservation biology uses estimates of animal abundances to examine effects of natural stochastic variables (e.g., El Niño Southern Oscillation, fires, floods) or anthropogenic activities (e.g., timber harvesting, mining, grazing, urbanization) on wildlife populations or communities (Aleixo 1999, Borges and Stouffer 1999, Buford and Capen 1999, Thompson et al. 1999). Many of these studies make management recommendations for entire communities based on their results, including recommendations for rare or infrequently encountered species. We considered species as rare (Rabinowitz 1981, Blackburn and Gaston 1997, Gaston 1997, Van Auken 1997) if they were encountered infrequently

Address for Dianna M. Queheillalt, Denise E. Taylor, and Michael R. Harris: Department of Biological Sciences, California State University, 6000 J Street, Sacramento, CA 95819-6077, USA; present address for Queheillalt: Department of Wildlife Ecology, College of Natural Sciences, University of Maine, 210 Nutting Hall, Orono, ME 04469, USA; e-mail: dmqueheillalt@hotmail.com. Address for James W. Cain III: School of Renewable Natural Resources, University of Arizona, Tucson, AZ 85721-0043, USA. Address for Michael L. Morrison: White Mountain Research Station, University of California, 3000 E. Line Street, Bishop, CA 93514-3637, USA. Address for Stacia L. Hoover: 3253 Cambridge Road, Cameron Park, CA 95682-9127, USA. Address for Natasha Tuatoo-Bartley: 7737 College Town Drive, #34, Sacramento, CA 95826-6077, USA. Address for Lourdes Rugge: 11276 E. Lime Kiln Road, Grass Valley, CA 95949-9719, USA. Address for Kirsten Christopherson: 9 CES/CEV, 6601 B Street, Beale AFB, CA 95903-1704, USA. Address for Miriam D. Hulst: 13419 Bentley Street, Waterford, CA 95386-9406, USA. Address for Heather L. Keough: Department of Fisheries and Wildlife, Utah State University, Logan, UT 84322-0001, USA.

by researchers. Therefore, our paper is not restricted to discussion of sensitive species (i.e., threatened and endangered). Species could be common in the wild but encountered infrequently due to their distribution or the sampling methodology. Such species are addressed here.

As we often encountered studies during the course of our research that excluded species from community-level analyses (e.g., correlations between species diversity and stand density) due to small sample sizes (or species rarity), we determined how frequently researchers did not use data collected on species in these analyses due to problems pertaining to small sample size.

Methods

To determine how frequently data on rare species were not analyzed or reported by researchers, we reviewed *Auk, Canadian Field Naturalist, Condor, Conservation Biology, Journal of Herpetology, Journal of Mammalogy, Journal of Wildlife Management, Southwestern Naturalist,* and *Wildlife Society Bulletin* from 1987–2001; Ecology from 1983–2001; and Ecological Applications from 1991–2001 for original articles that involved measures or indices of abundance for some or all of the species sampled. We selected journals based on their representation of a wide breadth of subjects in wildlife management.

We further reviewed the articles to determine number of species sampled, number of species included in community-level analyses, number of species excluded from such analyses, and the reason, if given, for not including all species in all calculations. We then tallied the number of articles that used all abundance or density estimates in community-level analyses for all species encountered. We also tallied articles that excluded some of the species sampled due to small sample sizes and articles that excluded species for reasons other than sample size. For articles in which the reason for exclusion was not clearly stated, we assumed that exclusion was done for reasons other than insufficient sample size so that our estimates would be conservative. If species were excluded from analyses due to a predetermined reason (e.g., the author was interested only in studying breeding bird species and therefore did not examine habitat preferences of migrants), we considered the article to have included all species for the purposes of our review. We calculated percentage of species excluded from community-level analyses due to small sample sizes for each article. We also calculated mean, standard deviation, and range for the percentage of species excluded due to small sample sizes.

Table 1. Articles from wildlife journals we reviewed that emphasized wildlife community ecology and included or excluded species from community-level analyses.

Journal	Articles reviewed	All species included	Species excluded due to small sizes	Species excluded for unstated reasons	Species excluded for pre-determined reasons
Auk	12	3	6	1	2
Canadian Field Naturalist	14	14	0	0	0
Condor	28	19	3	1	5
Conservation Biology	21	12	4	4	1
Ecological Applications	17	3	13	0	1
Ecology	14	10	3	1	0
Journal of Herpetology	10	8	2	0	0
Journal of Mammalogy	24	13	8	0	3
Journal of Wildlife Management	13	6	6	0	1
Southwestern Naturalist	24	12	10	1	1
Wildlife Society Bulletin	13	9	4	0	0
Total	190	109	59	8	14

Results

Thirty-one percent ($n = 59$) of the articles we reviewed did not include all species in community-level analyses due to small sample sizes, 7.4% ($n = 14$) excluded species due to predetermined reasons, and 57.4% ($n = 109$) included all species in their calculations (Table 1). The remaining articles (4.2%, $n = 8$) excluded species from their calculations for undisclosed reasons. *Ecological Applications* (76%) contained the highest percentage of articles excluding rare species. *Auk, Ecological Applications, Journal of Herpetology, Journal of Wildlife Management,* and *Southwestern Naturalist* had high (42–76%) numbers of articles excluding species from abundance or density calculations. Of the 59 articles excluding species due to small

Table 2. Mean ($\bar{x}$), standard deviation (SD), and range for percentage of species excluded from articles not including all abundance and density estimates in the analyses.

Journal	$\bar{x}$	SD	Range	*n*
Auk	28	17.1	11–57	5
Condor	58	33.2	34–81	2
Conservation Biology	68	4.6	65–71	2
Ecological Applications	51	23.5	21–93	12
Ecology	44	9.5	35–54	3
Journal of Herpetology	49	22.6	33–65	2
Journal of Mammalogy	60	20.8	15–78	8
Journal of Wildlife Management	58	18.9	35–80	6
Southwestern Naturalist	44	24.7	18–88	10
Wildlife Society Bulletin	55	20.2	32–61	3
All journals	50	21.9	11–93	53

sample sizes and stating how many species were sampled or excluded, an average of 50% of the species were excluded from abundance or density calculations (Table 2).

Discussion

Due to the large number of rare species in plant and wildlife communities, most communities adhere to log-normal species abundance distributions, in which a small number of species are common, only a few species reside in intermediate to low numbers, and most are uncommon (Rosenberg et al. 1995, Van Auken 1997, Harte et al. 1999, Maina and Howe 2000). Frequently used sampling designs, such as simple random sampling, stratified random sampling, and systematic random sampling, are useful when examining common species but ineffective when applied to infrequently encountered species (Thompson 1992, Thompson et al. 1998, Morrison et al. 2001). Such sampling designs return numerous zero counts and decrease the accuracy of studies using these designs (Thompson et al. 1998).

Nearly a third of the articles we reviewed excluded species from community-level analyses due to low sample sizes. Of these articles, half of the species were excluded from statistical analyses. Thus, these articles made management recommendations for entire communities or areas based on less than half of the species present. Because rare species are often limited in their geographic range or low in abundance, management decisions should not be made for a community based on the more frequently encountered species (Simberloff and Gotelli 1984, Thomas and Mallorie 1985, Bolger et al. 1991, Given and Norton 1993).

The exclusion of species due to low detection rates leads to the erroneous inflation of relative abundance calculations of included species. In many instances this may be trivial, but in instances of special-status species (e.g., threatened or endangered under state or federal law), elevated abundance estimates may lead to the conclusion that a species is prevalent in sufficient numbers when in fact its abundance is actually low.

If the object of a study is to compare relative abundances over successive years, trends may appear for a species that are due to the number of species excluded from abundance calculations rather than true biological trends. For example, if 1 year of data resulted in 5 species being excluded from calculations of relative abundance followed by 2 species being excluded from calculations for the subsequent year, a species encountered may be falsely thought to have increased in numbers as an artifact of statistical calculations rather than reflecting a true biological trend.

Management implications

Because rare species often are spatially clumped, we recommend using one of the forms of adaptive sampling methods, such as adaptive cluster sampling design, strip adaptive cluster sampling, or stratified adaptive cluster sampling (Thompson 1992). Adaptive cluster sampling is a two-stage sampling design in which initial sampling plots are randomly selected and monitored. Any of the initial plots containing animals are selected to have all adjacent plots monitored. This process continues until adjacent plots no longer contain animals of interest (Thompson 1992, Krebs 1999, Morrison et al. 2001). Strip adaptive cluster sampling and stratified adaptive cluster sampling are modified forms of adaptive cluster sampling (for further details see Thompson 1990; 1991*a*, *b*; 1992). Because adaptive sampling methods increase the probability of encountering clumped species and thus often increase sample sizes, they are an adequate choice of sampling methodology for rare species.

In addition, a higher number of smaller sampling units are preferable over fewer larger units in detecting rare species (Green and Young 1993) because the probability of initially detecting a rare species is increased. Adjacent sampling units can then be chosen as described above for adaptive cluster sampling to insure the maximum number of detections for rare species.

Samples from highly variable populations, such as those from communities with a large number of rare species, often have low statistical power (Peterman 1990). However, there are options for statistical analyses with small sample sizes. Contingent upon the specific situation and type of data being used, nonparametric tests can be employed or data transformed to allow the use of parametric tests when working with small sample sizes.

When assessing abundances and densities in wildlife communities, it is imperative to make every effort possible to include all species detected in community-level analyses to more accurately assess the effects of natural variables or anthropogenic activities on a community. This can be accomplished by altering the sampling design to an adaptive cluster method and increasing the number of sampling units. Should there still be small sample sizes for some species, statistical alternatives exist to excluding rare species from calculations of abundance or density estimates. These include the use of indices rather than absolutes, transforming data when possible prior to analysis, or the use of nonparametric tests when possible. By reducing the number of species excluded due to small sample sizes, researchers can make more accurate assessments of the effects of natural and anthropogenic factors on wildlife communities.

Acknowledgments. We thank P. R. Krausmann and K. S. Smallwood for their comments on earlier drafts of this manuscript.

Literature cited

Aleixo, A. 1999. Effects of selective logging on a bird community in the Brazilian Atlantic forest. Condor 101:537–548.

Blackburn, T. M., and K. J. Gaston. 1997. Who is rare? Artifacts and complexities of rarity determination. Pages 48–58 *in* W.E. Kunin and K.J. Gaston, editors. The biology of rarity. Chapman and Hall, London, United Kingdom.

Bolger, D. T., A. C. Alberts, and M. E. Soulé. 1991. Occurrence patterns of bird species in habitat fragments: sampling, extinction, and nested species subsets. American Naturalist 137:155–166.

Borges, S. H., and P. C. Stouffer. 1999. Bird communities in two types of anthropogenic successional vegetation in central Amazonia. Condor 101:529–536.

Buford, E., and D. Capen. 1999. Abundance and productivity of forest songbirds in a managed, unfragmented landscape in Vermont. Journal of Wildlife Management 63:180–188.

Gaston, K. J. 1997. What is rarity? Pages 30–41 *in* W. E. Kunin and K. J. Gaston, editors. The biology of rarity. Chapman and Hall, London, United Kingdom.

Given, D. R., and D. A. Norton. 1993. A multivariate approach to assessing threat and for priority setting. Biological Conservation 64:57–66.

Green, R. H., and R. C. Young. 1993. Sampling to detect rare species. Ecological Applications 3:351–356.

Harte, J., A. Kinzig, and J. Green. 1999. Self-similarity in the distribution and abundance of species. Science 284:334–336.

Krebs, C. J. 1999. Ecological methodology. Second edition. Cummings, Menlo Park, California, USA.

Maina, G. G., and H. F. Howe. 2000. Inherent rarity in community restoration. Conservation Biology 14:1335–1340.

Morrison, M. L., W. M. Block, M. D. Strickland, and W. L. Kendall. 2001. Wildlife study design. Springer-Verlag, New York, New York, USA.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Science 47:2–15.

Rabinowitz, D. 1981. Seven forms of rarity. Pages 205–217 *in* H. Synge, editor. The biological aspects of rare plant conservation. Wiley, Chichester, United Kingdom.

Rosenberg, D. K., W. S. Overton, and R. G. Anthony. 1995. Estimation of animal abundance when capture probabilities are low and heterogeneous. Journal of Wildlife Management 59: 52–261.

Simberloff, D., and N. Gotelli. 1984. Effects of insularization on plant species richness in the prairie forest ecotone. Biological Conservation 29:27–46.

Thomas, C. D., and H. C. Mallorie. 1985. Rarity, species richness and conservation: butterflies of the Atlas Mountains in Morocco. Biological Conservation 33:95–117.

Thompson, I. D., H. A. Hogan, and W. A. Montevecchi. 1999. Avian communities of mature balsam fir forest in Newfoundland: age-dependence and implications for timber harvesting. Condor 101:311–323.

Thompson, S. K. 1990. Adaptive cluster sampling. Journal of the American Statistical Association. 85:1050–1059.

Thompson, S. K. 1991*a*. Stratified adaptive cluster sampling. Biometrika 78:389–397.

Thompson, S. K. 1991*b*. Adaptive cluster sampling: designs with primary and secondary units. Biometrics 47:1103–1115.

Thompson, S. K. 1992. Sampling. Wiley, New York, New York, USA.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic, San Diego, California, USA.

Van Auken, O. W. 1997. Species rareness and commonness along spatial and temporal gradients. Southwestern Naturalist 42:369–374.

Associate Editor: *Euler*