J. Field Ornithol., 65(2):210-213

TESTING FOR RESOURCE USE AND SELECTION BY MARINE BIRDS: A COMMENT

NICHOLAS J. AEBISCHER AND PETER A. ROBERTSON
The Game Conservancy Trust
Fordingbridge, Hampshire SP6 1EF United Kingdom

Abstract.—The method recently advocated by Haney and Solow (1992: J. Field Ornithol. 63:43-52) for the analysis of resource use and selection does not recognize the underlying structure of the data, so is inappropriate. In particular, it ignores the unit-sum constraint, requires independence between observations, and does not allow for extra-binomial variability. Compositional analysis, which is tailored to the data structure, and provides power and flexibility in its application, is recommended instead.

PRUEBAS RELATIVAS AL USO Y SELECCIÓN DE RECURSOS POR AVES MARINAS: UN COMENTARIO

Sinopsis.—El método propuesto por Haney y Solow (1992: J. Field Ornithol. 63:43-52) para el análisis de la selección y utilización de recursos por aves marinas, es inapropiado debido a que éste no reconoce adecuadamente la estructura de los datos. En particular el mismo ignora las restricciones de la suma-unidades, requiere independencia entra observaciones y no da lugar a la variabilidad extra-binomial. Se recomienda un análisis composicional que se adapte a la estructura de los datos y provea de fuerza y flexibilidad en sus aplicaciones.

Data describing the proportional use of resources in relation to their availability are commonly collected by wildlife biologists. The analyses of such data are not necessarily straightforward, and it has been argued that many currently recommended statistical methods are inappropriate (Aebischer et al. 1993b).

The recent paper by Haney and Solow (1992) suggests using a technique for calculating simultaneous confidence intervals (Neu et al. 1974, clarified in Byers et al. 1984) to test marine resource preference based on count (enumeration) data. For each resource type i, the technique is to calculate the binomial standard error of the observed proportion of use p_i, then to evaluate the confidence interval using a Bonferroni z-statistic (Miller 1981:219), which adjusts for simultaneous estimation across all resource types. If the expected available proportion of a resource is outside the Bonferroni confidence interval, then that resource is identified as being used non-randomly.

Although this technique attempts to control the experiment-wise error rate, it is in fact inappropriate because it does not recognize the underlying biological and statistical structure of the data (Aebischer et al. 1993b). The problems are outlined below.

(1) The proportions describing use or availability sum to one over all k resource types; this is known as the unit-sum constraint (Aitchison 1986). It means that the measure of use p_i of resource type i depends on the use of the other k-1 resource types (mathematically, the multivariate response $\{p_1, \ldots, p_k\}$ is of dimension k-1, not k). The confidence-interval method ignores the constraint and proceeds as though the test for random

use of a given resource type were independent of the others. As well as calculating the Bonferroni z-statistic under the false assumption that $\{p_1, \ldots, p_k\}$ is of dimension k, this procedure can lead to a misinterpretation of resource selection. The problem is best seen in the case of just two resource types, with proportional use p_1 and p_2. The size of p_1 is inversely related to that of p_2 through the unit-sum constraint $p_1 + p_2 = 1$ ($p_1 = 1 - p_2$). The test comparing the use of resource type 1 with its availability is the same as the corresponding test for resource type 2, so that significant selection for type 1 automatically induces apparent significant avoidance of type 2.

(2) The conditions of binomial experiments that underpin the confidence-interval method (independence between individual birds, equal probability of occurrence within a resource type) are violated if the observations used to estimate the distribution of resource use are counts of seabirds that flock in a non-independent fashion (rafts of shearwaters *Puffinus* spp., socially-facilitated feeding flocks [Haney et al. 1992, Hoffman et al. 1981]). The effect is to inflate the sample size used by the confidence-interval method, leading to an underestimation of the confidence intervals and an increase in Type 1 error (rejection of a true null hypothesis).

(3) Data on resource use are usually collected in discrete "packets" (e.g., transects that each provide an estimate of seabird distribution, individual birds in the case of diet based on stomach contents). These "packets" are snapshots of the distribution of resource use under a range of conditions; the distributions thus sampled do not necessarily come from a common underlying multinomial distribution. The confidence-interval method pools the data from all the "packets," thereby ignoring the data structure. Such pooling is acceptable only if the assumption of a common underlying multinomial distribution is tested and found valid. Failure of the assumption leads to overdispersion, violation of the conditions of binomial experiments and, again, to an increase in Type 1 error.

(4) A similar problem arises when groups of birds behave differently, for instance when age or sex affects behaviour, and when the resource use data are pooled regardless of such differences. The confidence-interval method is not sufficiently flexible to accommodate such situations, nor does it allow testing for such effects. Log-linear modeling may be a suitable alternative (Heisey 1985) but, like the confidence-interval method, assumes a multinomial or product multinomial distribution of the observations (Everitt 1992).

In contrast, the technique of compositional analysis (Aitchison 1986) is tailor-made to analyze multivariate proportional data of the sort encountered in measurements of resource use and availability. It circumvents the unit-sum constraint by transformation of a set of k constrained proportions describing resource composition into a set of k-1 unconstrained values, which are then amenable to multivariate analysis of variance/covariance (MANOVA/MANCOVA). The range of situations and hypotheses encompassed by such linear models makes compositional analysis a versatile and powerful tool. Hypothesis testing is based on variability estimated from the data, rather than on the mean-variance relationship of binomial theory. Standard tests assume multivariate normality, but can be replaced by randomization tests (Manly 1991) if this assumption is false.

Using compositional analysis, the comparison of proportional resource use and availability becomes a multivariate analogue of the paired t-test. This is suitable for data in "packets," where the use and availability is known (and different) for each "packet." The full details are described in Aebischer et al. (1993b), and the practical computations in Aebischer and Robertson (1992). Differences in resource use according to, say, age, sex, weather or time of day, can be modeled and tested by MANOVA (Aebischer et al. 1993a). The computations are performed by most statistical computer packages, e.g., SYSTAT (Wilkinson 1990); a program implementing the randomization tests is available on demand from J. P. Carroll (Department of Biological and Environmental Sciences, California University of Pennsylvania, 250 University Avenue, California, Pennsylvania 15419).

Haney and Solow (1992) give two examples using the confidence-interval method. We describe below where the dangers lie in their approach, and how compositional analysis may be applied instead.

The first example is based on line transect observations of Short-tailed Shearwaters (*Puffinus tenuirostris*) in relation to four different oceanographic habitats. Each transect ("packet") provides an estimate of the distribution of resource use and availability over a given period. The confidence-interval approach not only ignores the unit-sum constraints (1) but also pools data across transects, thereby ignoring the variability between transects and the "packeted" nature of the data (3). By using birds as independent observations, the risk of inflated sample sizes is high (2), especially given the highly gregarious nature of non-breeding Short-tailed Shearwaters (Hoffman et al. 1981). Compositional analysis would treat each transect as a paired sample of shearwater distribution and habitat distribution made up of proportions describing habitat use and availability; it thereby respects the data structure. Moreover, it can adjust for covariables (such as weather, time of day or latitude) that may help explain inter-transect variability but that invalidate the assumption, made by the confidence-interval method, of a single underlying multinomial distribution common to all transects.

The second example considered the diet of the Crested Auklet (*Aethia cristatella*). It tested the hypothesis that auklets used all of six prey types equally, by comparing the observed proportions to 1/6. It must be pointed out that with no information on prey availability, testing this hypothesis is biologically meaningless. It is certainly not possible to "state confidently that Crested Auklets preferentially preyed upon *Thysanoessa* euphausiids" (Haney and Solow 1992:48) because *Thysanoessa* spp. could dominate in terms of availability. Setting aside the question of availability, we note that Haney and Solow (1992) used the number of stomachs rather than the number of prey items contained in the stomachs as their sample size; this approach is correct, as the "packets" of data are indeed the stomachs (3). In order to do so, the authors have classified the stomachs according to numerically dominant prey type, and ignored the detailed dietary composition of each stomach. This loss of information could have been avoided by the use of compositional analysis, which would have taken into account the unit-sum constraint at the same time (1).

In conclusion, the confidence-interval method advocated by Haney and Solow (1992) is inappropriate for analyzing resource use and selection. We strongly recommend the use of compositional analysis, which is specifically designed to model the structure of resource use data, and brings the versatility of MANOVA/MANCOVA to such analyses.

LITERATURE CITED

AEBISCHER, N. J., V. MARCSTRÖM, R. E. KENWARD, AND M. KARLBOM. 1993a. Survival and habitat utilisation: a case for compositional analysis. Pp. 343–353 *in* J.-D. Lebreton and P. M. North, eds. Marked individuals in the study of bird population. Birkhäuser Verlag, Basel, Switzerland.

———, AND P. A. ROBERTSON. 1992. Practical aspects of compositional analysis as applied to pheasant habitat utilization. Pp. 285–293, *in* I. G. Priede and S. M. Swift, eds. Wildlife telemetry: remote monitoring and tracking of animals. Ellis Horwood, New York, New York.

———, ———, AND R. E. KENWARD. 1993b. Compositional analysis of habitat use from animal radio-tracking data. Ecology 74:1313–1325.

AITCHISON, J. 1986. The statistical analysis of compositional data. Chapman and Hall, New York, New York. 416 pp.

BYERS, C. R., R. K. STEINHORST, AND P. R. KRAUSMAN. 1984. Clarification of a technique for analysis of utilization-availability data. J. Wildl. Manage. 48:1050–1053.

EVERITT, B. S. 1992. The analysis of contingency tables. 2nd edition. Chapman and Hall, New York, New York. 164 pp.

HANEY, J. C., K. M. FRISTRUP, AND D. S. LEE. 1992. Geometry of visual recruitment by seabirds to ephemeral foraging flocks. Ornis Scand. 23:49–62.

———, AND A. R. SOLOW. 1992. Testing for resource use and selection by marine birds. J. Field Ornithol. 63:43–52.

HEISEY, D. M. 1985. Analyzing selection experiments with log-linear models. Ecology 98:437–456.

HOFFMAN, W., D. HEINEMANN, AND J. A. WIENS. 1981. The ecology of seabird feeding flocks in Alaska. Auk 98:437–456.

MANLY, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, New York, New York. 281 pp.

MILLER, R. G. 1981. Simultaneous statistical inference. 2nd edition. Springer-Verlag, New York, New York. 299 pp.

NEU, C. W., C. R. BYERS, AND J. M. PEEK. 1974. A technique for analysis of utilization-availability data. J. Wildl. Manage. 38:541–545.

WILKINSON, L. 1990. SYSTAT: the system for statistics. Systat, Inc., Evanston, Illinois. 677 pp.

Received 16 Mar. 1993; accepted 6 Jul. 1993.