Misclassified Resource Selection: Compositional Analysis and Unused Habitat

Author(s): RALPH L. BINGHAM, LEONARD A. BRENNAN, BART M. BALLARD
Source: Journal of Wildlife Management, 71(4):1369-1374.
Published By: The Wildlife Society
https://doi.org/10.2193/2006-072
URL: http://www.bioone.org/doi/full/10.2193/2006-072

Techniques and Technology Note

Misclassified Resource Selection: Compositional Analysis and Unused Habitat

RALPH L. BINGHAM,[1] *Caesar Kleberg Wildlife Research Institute, Texas A&M University-Kingsville, Kingsville, TX 78363, USA*
LEONARD A. BRENNAN, *Caesar Kleberg Wildlife Research Institute, Texas A&M University-Kingsville, Kingsville, TX 78363, USA*
BART M. BALLARD, *Caesar Kleberg Wildlife Research Institute, Texas A&M University-Kingsville, Kingsville, TX 78363, USA*

ABSTRACT During the past decade, compositional analysis (CA) has been used widely in animal–habitat and resource selection studies. Despite this popularity, CA has not been tested for potential systematic biases such as incorrect identification of preferred resources. We used computer-simulated data based on known habitat use and availability parameters to assess the potential for CA to incorrectly identify preferred habitat use. We consider in particular the situation when available habitat categories not used by all animals are included in the resource selection analysis, with substitution of a relatively small value, such as 0.01, for each 0% utilization value. Progressively larger misclassification-error rates in preferred habitat use resulted from substituting progressively smaller positive values for each 0% utilization of a habitat category.

(JOURNAL OF WILDLIFE MANAGEMENT 71(4):1369–1374; 2007)

DOI: 10.2193/2006-072

KEY WORDS compositional analysis, known habitat use and availability parameters, misclassification error rate, Monte Carlo simulations, resource selection analyses.

Compositional analysis (CA) has been recommended as a preferred method of quantifying habitat use by free-ranging animals (Aebischer et al. 1993). This method has been used to analyze home range and habitat use by a wide variety of animals in many different places (Bingham and Brennan 2004). If study animals are not observed or detected in one or more available habitat categories analyzed using CA, a data modification is required so that the log-ratio representing use of a habitat category can be calculated. The procedure recommended by Aitchison (1986) and Aebischer et al. (1993) to solve this problem was to substitute an arbitrarily small positive value, such as 0.01, for each 0% utilization value for any animal. Aitchison (1986:270) realized this was a potential problem in CA when he noted, "Even after a compositional data set has been adjusted for zero. . .it should still be subjected to some form of sensitivity analysis." With respect to this potential problem in CA, Aebischer et al. (1993:1320) asserted, "Results therefore seem robust with respect to the choice of value to replace a 0% utilization of an available habitat type, provided that the value is less than existing nonzero values in either available or utilized compositions. This fits in with the rationale that 0% represents a use too low to be recorded, so should be replaced by a value distinctly less than the smallest nonzero value: an order of magnitude less is probably appropriate to most situations." This recommendation appears to have been interpreted as smaller is better by some researchers who have utilized CA with substitution values of 0.001 (Miller et al. 1999, Chamberlain et al. 2000) and 0.0001 (Hartke and Hepp 2004). Aebischer et al. (1993:1322) also stated, "The performance of the technique remains to be evaluated, as the only true yardstick is simulated data based on known parameters." In relation to the problem of 0% utilization of habitat categories in CA, Pendleton et al. (1998:290) noted, "Simulations can be used to assess the effect of varying values of the constant or threshold." These authors appear to assume CA is robust to this potential analytical weakness, or that such weakness can be overcome with randomized simulation analyses of empirical data.

Bingham and Brennan (2004) used simulation analyses to examine the potential of CA to produce Type I errors when arbitrarily small values were substituted for 0% use of available habitat categories. They found that, under these circumstances, Type I error rates ranged from 5% to 100%, and were highest (>50–100%) when the smaller (0.001) positive values were substituted for 0% utilization of habitat categories. Since such a great potential exists for Type I errors in CA under these circumstances, we thought it would be informative to conduct a companion study that examined the potential for misclassification of preferred habitat use (which can either be a Type I or Type II error) to occur in CA when arbitrarily small positive values are substituted for 0% utilization. Dasgupta and Alldredge (2002) also observed that CA resulted in high Type I error rates for simulations of habitats involving relatively low availabilities, although they give no indication as to what was done about the occurrence of 0% use.

Our objectives were to 1) use simulated data based on known habitat use and availability parameters to evaluate misclassification of preferred habitat-use error rates for CA when arbitrarily small positive values are substituted for 0% use of a habitat category and 2) compare these CA misclassification error rates for the same habitat configuration with the same relative utilization but with a sufficient number of observations to result in all animals being observed in all available habitats.

METHODS

The CA method uses multivariate analysis of variance (MANOVA) models to analyze log-ratios for comparison of

---

[1] *E-mail: ralph.bingham@tamuk.edu*

Table 1. Overall percentages of utilization for all animals in each simulation resulting from restricted randomization of use units to 5 habitat categories with fixed availabilities, where 20% of use units for each animal were randomly assigned to habitats 1–3 and the remaining 80% were assigned to habitats 4–5.

Habitat	Availability (%)	Utilization (%)
1	1	2
2	3	6
3	6	12
4	40	36
5	50	44

utilization and availability of habitats (Aebischer et al. 1993). Assume that D habitat categories are available and that an individual animal's proportional use is represented by the composition $U_1, U_2, \ldots, U_D$, where U_i is the proportion of observations in habitat of category i for $i = 1, 2, \ldots, D$. The sum of the U_is is 1. Similarly, assume the proportions of available habitat for the same animal are represented by the composition $A_1, A_2, \ldots, A_D$. For any component x_j of a composition, the log-ratio transformation $y_i = \ln(x_i/x_j)$; ($i = 1, 2, \ldots, \mathrm{D}, i \neq j$) results in new y_is that are linearly independent (i.e., "a one-to-one map of a point on a D-dimensional simplex to a point in full ($D - 1$)-dimensional space" [Aebischer et al. 1993:1315]). The $D - 1$ differences $d_{iD} = \ln(U_i/U_D) - \ln(A_i/A_D) = \ln(U_i/A_i) - \ln(U_D/A_D)$ for $i = 1, 2, \ldots, D - 1$ are calculated for each individual animal. If random selection occurs, the mean value of d_{iD} over all animals should be zero for all i (as noted by Aebescher et al. (1993:1315) who emphasized that in such a case "$d \equiv 0$"). The MANOVA test for selection (i.e., nonrandom use), therefore, consists of testing simultaneously over all habitats for $i = 1, 2, \ldots, D - 1$ whether the vector of mean values of d_{iD} is significantly different from the zero vector. We interpreted a significant value of a test statistic, such as Wilks' lambda, as indicating that nonrandom use has occurred.

The experimental units for CA are the animals, which we assumed independent because CA is rooted in MANOVA. Assuming we determined nonrandom use, the next step was to find where use deviated from random, and which habitat category was used more than expected (obs relative frequency is more than available relative frequency) relative to another habitat category. The expression for the pair-wise differences of log-ratios $d_{ij} = \ln(U_i/U_j) - \ln(A_i/A_j)$ for $i \neq j$ is equivalent to $\ln(\mathrm{U}_i/\mathrm{A}_i) - \ln(\mathrm{U}_j/\mathrm{A}_j) = \ln[(\mathrm{U}_i/\mathrm{A}_i)/(\mathrm{U}_j/\mathrm{A}_j)]$ which is the logarithm of the ratio of preferences. Therefore, because $\mathrm{d}_{ij} = 0$ only if $\mathrm{U}_i/\mathrm{A}_i = \mathrm{U}_j/\mathrm{A}_j$, paired t-tests can then be used to determine whether a habitat category is used more than expected relative to each of the other habitat categories (Aebischer et al. 1993; i.e., whether a habitat category is preferred over another habitat category). Given that we found significant nonrandom use, Aebischer et al. (1993:1316) recommend "... staying with standard significance levels for t rather than, say, Bonferroni levels, by analogy with the protected least-significant-difference procedure."

Our hypothetical study area was a 10 × 10 square with each of the resulting 100 square units labeled as a particular habitat category 1–5. We used the number of squares assigned to each habitat category as the availability percentage. We assigned 1 square, 3 squares, 6 squares, 40 squares, and 50 squares to habitat categories 1–5, respectively. We chose these availability percentages to be similar to those reported for the ring-necked pheasant (*Phasianus colchicus*) example of Aebischer et al. (1993). Also, we chose the smaller (e.g., 1%, 3%, and 6%) availability values to ensure the occurrence of 0% utilization for some of the animals when modeled with only 30 observations for each animal.

We simulated 3 different animal-observation combinations: 30 observations per animal for 30 and 100 animals, and 500 observations per animal for 30 animals. We randomly assigned 20% of the observations for each animal to the 3 habitat categories with 1%, 3%, and 6% availability, and randomly assigned the remaining 80% of the observations to the 2 habitat categories with 40% and 50% availability. We restricted randomization to obtain overall relative utilization of each of the habitat categories with 1%, 3%, and 6% availability that was twice their relative availability, thus making these categories preferred (Neu et al. 1974). Overall relative utilization was less than relative availability for each of the habitat categories with 40% and 50% availability, thus making these categories avoided (Table 1). Thus, a Type II error would occur if we did not find a habitat category with 1%, 3%, or 6% availability to be preferred over a habitat category with 40% or 50% availability. Additionally, we designed the randomization scheme to indicate no preference among habitat categories with 1%, 3%, and 6% availability and no preference between the remaining habitat categories with 40% and 50% availability. This represented the Type I error aspects of our analyses. This results in each of the habitat categories with 1%, 3%, and 6% availability being labeled as equally preferred, and each of the remaining habitat categories with 40% and 50% availability, labeled as equally avoided, or not preferred. We tested for nonrandom utilization using 1,000 Monte-Carlo simulations for each of the 3 animal-observation combinations for the 5 habitat categories, resulting in 3,000 simulations. For the simulations that we ran for 30 animals and 100 animals with 30 observations per animal (i.e., 1,000 simulations each), we made 7 different substitutions of 0.001% to 0.9% for each 0% use, resulting in a total of 15,000 Wilks' lambda tests for nonrandom use. We simulated random points within our study area for each animal using the SAS (SAS Institute, Inc., Cary, NC) pseudorandom number function RANUNI for each of the 3 animal-observation conditions.

We determined the number of correct preferential classifications by comparing each pair of habitats and counting how many simulations indicated a correct preference ($P < 0.05$) or no preference between the 2

Table 2. Percentage of simulations indicating significant nonrandom use, where a habitat was found to be preferred over another habitat category using compositional analysis (CA). (30 animals, 30 observations/animal).

	Substitution value for 0% use[a]							No 0% use[b]	Exp values
	0.001	0.01	0.1	0.3	0.5	0.7	0.9		
Preference of:									
1% over 3%[c]	0	0	0	0	0	2	6	1	0
3% over 1%	74	66	39	17	7	2	1	5	0
1% over 6%	0	0	0	0	0	0	1	0	0
6% over 1%	100	99	84	51	29	12	4	7	0
1% over 40%	0	0	0	9	53	96	100	100	100
40% over 1%	97	82	20	0	0	0	0	0	0
1% over 50%	0	0	0	8	54	97	100	100	100
50% over 1%	97	83	21	0	0	0	0	0	0
3% over 6%	0	0	0	0	0	0	0	2	0
6% over 3%	26	23	16	13	10	9	8	3	0
3% over 40%	4	10	34	63	82	93	97	100	100
40% over 3%	1	0	0	0	0	0	0	0	0
3% over 50%	4	11	33	64	82	94	98	100	100
50% over 3%	1	0	0	0	0	0	0	0	0
6% over 40%	92	98	100	100	100	100	100	100	100
40% over 6%	0	0	0	0	0	0	0	0	0
6% over 50%	92	97	100	100	100	100	100	100	100
50% over 6%	0	0	0	0	0	0	0	0	0
40% over 50%	3	3	3	3	3	3	3	2	0
50% over 40%	3	3	3	3	3	3	3	2	0

[a] The range of substitution values (0.001–0.9%) for 0% use of a habitat, where 20% of observations for each animal were randomly chosen from habitats with 1%, 3%, and 6% availability and 80% were randomly chosen from habitats with 40% and 50% availability in a hypothetical study area, for 30 animals with 30 observations per animal, compared to compositional analysis results for 30 animals with 500 observations per animal where no 0% use occurred. Expected values for the simulations are provided in the far right column.

[b] Results for 30 animals with 500 observations/animal.

[c] Values are % availability of different habitat categories. Note: The no. of simulations out of 1,000 indicating significant ($P < 0.05$) nonrandom use for substitutions 0.001, 0.01, and 0.1 were 961, 949, and 1,000, respectively. All 1,000 simulations for each substitution value >0.1 and no 0% use indicated significant ($P < 0.001$) nonrandom use, all using Wilks' lambda.

habitat categories. We did this by comparing the results of the simulation analyses with various substitutions for 0% habitat use with the results of simulations based on habitat categories where animals occupied all patches. We considered a habitat category with overall relative utilization greater than its relative availability found to be preferred ($P < 0.05$) over a habitat category with overall relative utilization less than its relative availability a correct classification. We also considered any 2 habitats, both with their overall relative utilization greater than their relative availability (habitats with 1%, 3%, and 6% availability) or both with their overall relative utilization less than their relative availability (habitats with 40% and 50% availability) where no preference ($P > 0.05$) was found between them correctly classified. We counted all other cases as incorrect classifications. We performed all statistical analyses with SAS version 8.1.

RESULTS

All 8,000 Wilks' lambda tests for 100 animals with 30 observations per animal (with 7 substitution values for 0% use) and 30 animals with 500 observations per animal correctly indicated significant ($P < 0.001$) nonrandom use of habitat categories. This was also the case for the 4,000 tests with substitution values >0.1 for 0% use for 30 animals with 30 observations per animal. For each of the substitution values ≤ 0.1 for 0% use for 30 animals with 30 observations per animal, ≥ 949 of the 1,000 (95%) Wilks' lambda tests correctly indicated significant ($P < 0.05$) nonrandom use (Tables 2, 3).

We did not observe inflated misclassification error rates for the simulated animal-observation combination with 500 observations per animal for 30 animals, where all animals were observed in all habitat categories for each simulation (second from right columns in Tables 2 and 3 differ only slightly from expected values). On the other hand, the CA method produced widely varying misclassification error rates for the other 2 simulated animal-observation combinations (30 observations/animal for 30 animals and 100 animals) where 0% utilization occurred for some animals in the habitats with 1%, 3%, and 6% availabilities (Tables 2–4). The misclassification error rate was dependent on the value (ranging from 0.001 to 0.9) substituted for 0% utilization by any animal. For 30 observations per animal with 30 animals, the pattern of misclassification error occurrence was similar to that for 100 animals. Misclassification error rates were generally inflated and highest when we substituted the smallest values (e.g., 0.001, 0.01) for 0% use (Tables 2, 3). The habitat categories with 1%, 3%, and 6% of available area contained all of the 0% use values for the 2,000 simulations of 30 animals and 100 animals with 30 observations per animal (Table 4).

We found the degree of misclassification was a function of

Table 3. Percentage of 1,000 simulations indicating significant nonrandom use, where a habitat category was found to be preferred another habitat category using compositional analysis (CA). (100 animals, 30 observations/animal)

	Substitution value for 0% use[a]							No 0% use[b]	Exp values
	0.001	0.01	0.1	0.3	0.5	0.7	0.9		
Preference of:									
1% over 3%[c]	0	0	0	0	0	2	14	1	0
3% over 1%	100	99	88	44	15	3	0	5	0
1% over 6%	0	0	0	0	0	0	1	0	0
6% over 1%	100	100	100	97	72	32	5	7	0
1% over 40%	0	0	0	22	98	100	100	100	100
40% over 1%	100	100	61	0	0	0	0	0	0
1% over 50%	0	0	0	19	98	100	100	100	100
50% over 1%	100	100	62	0	0	0	0	0	0
3% over 6%	0	0	0	0	0	0	0	2	0
6% over 3%	88	82	64	47	37	28	22	3	0
3% over 40%	1	10	71	99	100	100	100	100	100
40% over 3%	4	0	0	0	0	0	0	0	0
3% over 50%	1	10	71	99	100	100	100	100	100
50% over 3%	4	0	0	0	0	0	0	0	0
6% over 40%	100	100	100	100	100	100	100	100	100
40% over 6%	0	0	0	0	0	0	0	0	0
6% over 50%	100	100	100	100	100	100	100	100	100
50% over 6%	0	0	0	0	0	0	0	0	0
40% over 50%	1	1	1	1	1	1	1	2	0
50% over 40%	4	4	4	4	4	4	4	2	0

[a] The range of substitution values (0.001–0.9%) for 0% use of habitat, where 20% of observations for each animal were randomly chosen from habitats with 1%, 3%, and 6% availability and 80% were randomly chosen from habitats with 40% and 50% availability in a hypothetical study area, for 100 animals with 30 observations per animal, compared to CA results for 30 animals with 500 observations per animal where no 0% use occurred. Expected values for the simulations are provided in the far right column.

[b] Results for 30 animals with 500 observations/animal.

[c] Values are % availability of different habitat categories. Note: All 1,000 simulations for each substitution value and no 0% use indicated significant ($P <$ 0.001) nonrandom use using Wilks' lambda.

the value substituted for 0% use. A substitution value >0.7% minimized the misclassification error rate for all pairs of habitat categories for the 2 animal-observation combinations where 0% utilization occurred (30 observations/animal for 30 animals and 100 animals) because of the particular way observations were modeled. When we used substitution values <0.3%, the number of simulations that incorrectly identified a habitat category as preferred ($P <$ 0.05) over another habitat category increased whenever the habitat category with 1% availability was compared to each of the other 4 habitat categories (Tables 2, 3). Consequently, the number of simulations that correctly identified the habitat category with 1% availability as preferred ($P <$ 0.05) over a habitat category with 40% or 50% availability decreased to zero for substitution values <0.3%.

Table 4. Average number of animals with 0% use for 1,000 simulations of 3 animal-observation conditions modeled with 20% of observations for each animal randomly assigned to habitats with 1%, 3%, and 6% availability and 80% randomly assigned to habitats with 40% and 50% availability in a hypothetical study area consisting of 5 habitat types.

No. animals	No. observations/ animal	Habitat availability (%)				
		1	3	6	40	50
30	30	16.0	3.6	0.1	[a]	[a]
100	30	53.0	11.7	0.4		

[a] No 0% use occurred for simulated animals in the 40% and 50% availability categories and all availability categories for 30 animals with 500 observations/animal.

As we made the substitution value closer to 0%, the CA misclassification error rate tended toward 100% for some pairs of habitat categories (Tables 2, 3). In particular, using a substitution value of 0.001% incorrectly indicated that each of the 4 habitat categories with 3%, 6%, 40%, and 50% availabilities was preferred ($P < 0.05$) over the habitat with 1% availability for 74–100% of the 961 simulations that were determined to indicate significant ($P < 0.05$) nonrandom use for 30 animals, and for 100% of the 1,000 simulations for 100 animals, with 30 observations per animal in both cases. This was contrary to the way we modeled observations for the habitat categories (Tables 2, 3). Also, the habitat with 1% availability was correctly identified as preferred ($P < 0.05$) over each of the 2 habitats with 40% and 50% availabilities only for substitution values >0.5% for all 1,000 simulations of 100 animals with 30 observations per animal (Tables 2, 3).

Similarly, using a substitution value of 0.001%, the habitat with 3% availability was correctly identified as preferred ($P < 0.05$) over each of the habitats with 40% and 50% availabilities for only 4% of the 961 simulations that were determined to indicate significant ($P < 0.05$) nonrandom use for 30 animals with 30 observations per animal, and for only 1% of the 1,000 simulations for 100 animals with 30 observations per animal (Tables 2, 3). On the other hand, each of the habitat categories with the 1% and 3% availabilities was correctly identified as preferred ($P < 0.05$) over each of the habitats with 40% and 50% availabilities in all 1,000 simulations of 30 animals with 500 observations per animal where no 0% use occurred. Further, a maximum of 7% of 1,000 simulations incorrectly indicated a preference ($P < 0.05$) between any 2 of the 3 habitat categories with the 1%, 3%, and 6% availabilities or between the 2 habitats with 40% and 50% availabilities using 30 animals with 500 observations per animal where no 0% use occurred (Tables 2, 3).

In general, increasing the number of animals (experimental units) from 30 to 100 with 30 observations per animal increased the misclassification error rate for all of the substitution values <0.9% where the habitat with 3% availability was found to be incorrectly identified as preferred over the habitat with 1% availability where the majority of cases with 0% use occurred (Tables 2–4). This same general trend occurred for incorrect preference of all habitats over the habitat with 1% availability and the incorrect preference of the habitat with 6% availability over the habitat with 3% availability. The counterintuitive aspect of this result must be a function of a systematic error associated with the substitution of an arbitrary nonzero value for each occurrence of 0% utilization of a habitat category. That is, an increase in the number of animals for a fixed number of observations per animal resulted in a higher frequency of 0% utilization for the 3 habitat categories with 1%, 3%, and 6% availability (Table 4).

All percentages in Tables 2 and 3 are based on the parametric *P*-values for both Wilks' lambda and the corresponding *t*-tests. A check of the corresponding randomized *P*-values for the relatively few marginally significant ($0.01 < P < 0.08$) cases did not indicate any change in the observed trends. Some randomized *P*-values were found to be slightly more than the corresponding parametric *P*-value, some slightly less, and in general no randomized *P*-value was observed to differ from the corresponding parametric *P*-value by >50% of the parametric *P*-value.

DISCUSSION

Whereas the CA statistical method did not indicate inflated misclassification error rates for resource selection analyses with relatively small habitat availabilities where no 0% utilization was observed for any animal, this was not the case when analyses included animals with 0% utilization of some habitat categories. When we used CA in these instances, we found the misclassification error rate to be minimized for all pairs of habitat categories for a substitution value greater than 0.7%, and approached 100% for some pairs of habitat categories using substitution values <0.3%, contrary to the robustness claimed by Aebischer et al. (1993). This particular pattern occurred because of the restricted random assignment of observations to the 5 habitat categories, where 20% of the observations for each animal were randomly assigned to the 3 habitats with smallest (1%, 3%, and 6%) availabilities (making them equally preferred) and the remaining 80% of the observations for each animal were randomly assigned to the 2 habitats with largest (40% and 50%) availabilities (making them equally not preferred or avoided). In general, such an apparent systematic bias is unacceptable for any analytical technique.

Whether some past studies that used CA to analyze habitat use of wild animals suffer from such a systematic bias is unknown, because authors seldom report how they dealt with the problem of 0% observations in an available habitat category. There are examples in the literature where an arbitrarily small substitution value such as 0.001 is used for habitat categories that received 0% use (Chamberlain 1999, Miller et al. 2001). Recently, Hartke and Hepp (2004) used an even smaller value (0.00001) to substitute for 0% use of a habitat category. Apparently, it has been assumed that smaller is better for substitution values to replace 0% use of habitat categories in CA. However, the smaller the substitution value, the higher the error rate both in some Type I (Dasgupta and Alldredge 2002, Bingham and Brennan 2004) and misclassification situations with CA. As the argument of a logarithmic function approaches zero, the value of the function approaches negative infinity. This may be a fundamental reason why CA has great potential to produce misleading results when log values are derived utilizing arbitrarily small positive values substituted for 0% use of habitat categories. As smaller substitution values are used for 0% utilization in CA, the opportunity for Type I and II errors can increase dramatically.

In relation to the substitution of a small positive value for 0% utilization of a habitat category, Pendleton et al. (1998:290) asserted "In practice, the choice of value often has relatively little effect on the qualitative test results. . ." Contrary to this assertion, our simulations indicate a huge potential exists for CA to introduce a predominant systematic bias in the form of a high misclassification error rate, when smaller positive values are substituted for 0% observations in an available habitat category. Possible misclassification error rates for CA should be considered unacceptable using the relatively smaller substitution values of 0.001 and 0.01 for 0% use. The latter value (0.01) was recommended and used by Aebischer et al. (1993) in their examples.

Misclassification error rates for the CA method became larger when more animals were modeled for the same number of observations per animal. This is, of course, the opposite of what is usually recommended for any statistical test using animals as experimental units. For the habitat condition examined with relatively small availabilities, and accompanying 0% utilization by some animals, CA should not be considered a satisfactory method to assess resource selection. A problem with CA is caused by its possibly large misclassification error rate dependent on the choice of a relatively small substitution value for 0% utilization, and also its possibly large Type I error rate that has been shown to be possible regardless of what substitution value is used for 0% utilization (Bingham and Brennan 2004).

We modeled scenarios where all 0% use observations occurred in habitat types with smaller availabilities due to the limitations of the restricted randomization methods that we employed. It is obvious that many other types of scenarios are possible, such as different numbers of habitats considered or the occurrence of 0% use in habitats with greater availabilities. Thus, simulations considering other scenarios and comparisons with other methods are still needed before any general advice can be given about what is the best way to apply CA when 0% use is observed. However, we clearly demonstrated that there is great potential for misclassification errors when 0% use occurs when using CA. Therefore, authors should be aware of these potential problems and practice caution when analyzing resource selection with CA or when evaluating other studies that have employed this method.

MANAGEMENT IMPLICATIONS

If CA is to be used for statistical analyses of resource selection, research or management experiments should be designed so the number of observations per radiomarked animal is sufficient to obtain at least a few observations in all available habitat categories. This may be possible in light of new technology such as satellite-tracking transmitters, or use of automated telemetry systems using fixed towers and data loggers. These methods can generate huge numbers of data points, compared to traditional methods where personnel manually collect 1–2 observations per radiomarked animal per day. An obvious remedy is to reclassify habitat categories so that no habitat categories with 0% use are included in CA (Aitchison 1986), or omit animals from CA that have not used some habitat categories as done by Aebischer et al. (1993) for one of their examples.

ACKNOWLEDGMENTS

We thank the Caesar Kleberg Wildlife Research Institute for supporting us in researching this particular subject. We also thank R. DeYoung and D. Hewitt for their helpful comments on the manuscript. This is publication number 05–116 of the Caesar Kleberg Wildlife Research Institute.

LITERATURE CITED

Aebischer, N. J., P. A. Robertson, and R. E. Kenward. 1993. 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, London, United Kingdom.

Bingham, R. L., and L. A. Brennan. 2004. Comparison of Type I error rates for statistical analyses of resource selection. Journal of Wildlife Management 68:206–212.

Chamberlain, M. J., B. D. Leopold, and L. W. Burger, Jr. 2000. Characteristics of roost sites of adult wild turkey females. Journal of Wildlife Management 64:1025–1032.

Dasgupta, N., and J. R. Alldredge. 2002. A single step method for identifying individual resources. Journal of Agricultural, Biological and Environmental Statistics 7:208–221.

Hartke, K. M., and G. R. Hepp. 2004. Habitat use and preferences of breeding female wood ducks. Journal of Wildlife Management 68:84–93.

Miller, D. A., G. A. Hurst, and B. D. Leopold. 1999. Habitat use of eastern wild turkeys in central Mississippi. Journal of Wildlife Management 63:210–222.

Neu, C. W., C. R. Beyers, and J. M. Peek. 1974. A technique for analysis of utilization-availability data. Journal of Wildlife Management 38:541–545.

Pendleton, G. W., K. Titus, E. DeGayner, C. J. Flatten, and R. E. Lowell. 1998. Compositional analysis and GIS for study of habitat selection by goshawks in southeast Alaska. Journal of Agricultural, Biological, and Environmental Statistics 3:280–295.

Associate Editor: Alpizar-Java.