

Comments

Ecology, 86(11), 2005, pp. 3119–3125
© 2005 by the Ecological Society of America

A COMPARISON OF DISTANCE-BASED AND CLASSIFICATION-BASED ANALYSES OF HABITAT USE: COMMENT

Christian Dussault,^{1,2} Jean-Pierre Ouellet,³ and Réhaume Courtois²

A method using Euclidean distances between observations and habitat types was recently proposed as an alternative to the classification-based methods to study habitat selection and is referred to as the distance-based approach (DA; Conner et al. 2003). Several potential advantages of DA over classification-based approaches have been put forward. First, distance to habitat patches is affected by patch size and shape, which may in turn influence habitat preference of animals (Helzer and Jelinski 1999). Also, DA theoretically allows detection of edge preferences, and results obtained with this method are less likely to be influenced by location error that can lead to misclassification when using a classification approach such as compositional analysis (CA; Conner et al. 2003). Finally, results obtained with DA should not be affected by the inclusion of unused but available habitat types in the analysis (Bingham and Brennan 2004).

In spite of potential advantages, we suggest that the DA method yields results that are difficult to interpret. It is possible, for example, to detect selection of a particular habitat type even when animals were never located within that habitat type, meaning that a habitat does not need to be used to be preferred (see Fig. 1 in Conner et al. 2003). Using this approach with a species that strongly prefers recently burned areas, for example, could lead to the conclusion that habitat types surrounding burned areas are also preferred, which could

Manuscript received 3 February 2005; revised 15 April 2005; accepted 18 April 2005. Corresponding Editor: A. M. Ellison.

¹ Département de Biologie, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, PQ, G5L 3A1 Canada. E-mail: christian.dussault@fapaq.gouv.qc.ca

² Direction de la Recherche sur la Faune, Faune Québec, 675 Boulevard René-Lévesque Est, Québec, PQ, G1R 5V7 Canada.

³ Département de Biologie, Centre d'Études Nordiques, Université du Québec à Rimouski, 300 Allée des Ursulines, Rimouski, PQ, G5L 3A1 Canada.

be incorrect. Also, we question the applicability of DA to situations where habitat types do not have similar patch sizes. If animals use a relatively large (or small) habitat patch with no regard to edge, then we expect mean distances to surrounding habitat types to be greater (or less) in that patch than in other sized patches. DA is thus likely not appropriate in landscapes where a habitat type systematically occurs in larger (or smaller) habitat patches relative to other habitat types. Clearly, the ability of DA results to correctly reflect resource preference under varying landscape structures needs to be assessed before relying on the method.

In this study, our objective was to evaluate the robustness of DA under varying habitat configurations and patch sizes. We do not question the validity or mathematical properties of DA, but we expected difficulties in interpreting DA results. We adopted exactly the same approach as Conner et al. (2003) and compared results obtained with DA and classification-based approaches using the same database. Comparisons were initially conducted using simulated data and then again using empirical data collected on moose (*Alces alces*). Similar to the terminology of Conner et al. (2003), we consider habitat selection to occur when habitat utilization differs from that expected by a null model. We define habitat types that are used more or less frequently than expected compared to availability as “selected” or “avoided,” respectively (Garshelis 2000). Finally, we define the habitat type that was used most relative to expected values as the “most preferred” and also the habitat type that was used least relative to expected values as the “least preferred.”

Analyses of habitat selection

All analyses of habitat selection compared use of habitat types, as measured by animal locations, to availability of those habitat types within individual home ranges, i.e., third-order selection of Johnson (1980). DA, as proposed by Conner et al. (2003), compares distances between animal locations and the nearest patch of each habitat type to distances obtained using random locations. Animals that are found to be closer to a habitat type than expected based on random locations are assumed to prefer that habitat type. Similar to Conner et al. (2003), results of DA were compared to those of CA, a classification-based approach (Aebischer et al. 1993). DA and CA both use individuals instead of individual radiolocations as the sampling unit and allow the assessment of the role of independent variables in selection patterns.

Simulations

Our hypothetical study area consisted of a 1600-km² block that was composed of eight different habitat types

TABLE 1. Patch size (mean \pm SE) and spatial association between habitat types in the moose study area.

Habitat	Percentage of edge shared with:						Total edge density (m/ha)	Patch size (ha)
	MI10	MI30	MI50	C30	DT50	OTHER		
MI10		5.7	35.4	55.2	0.2	3.5	25.6	18.8 \pm 3.4
MI50	30.9	6.2		51.3	1.4	10.2	29.2	30.2 \pm 4.2
MI30	14.3		17.9	54.1	2.3	11.5	10.2	20.3 \pm 2.0
DT50	4.8	21.0	39.0	1.0		34.3	1.2	125.9 \pm 50.9
OTHER	6.4	8.5	21.4	61.1	2.6		13.9	19.2 \pm 5.2
C30	32.7	12.8	34.8		<0.1	19.6	43.0	54.1 \pm 16.9

Notes: Spatial association was assessed using proportion of edge that each habitat type shares with five other habitat types. See *Empirical data* for descriptions of habitat types.

occurring in equal availability (12.5%), which is representative of most wildlife-habitat relationship studies (Bingham and Brennan 2004). Habitat patches were square in shape and habitat types were designated habitat A through H. Six different habitat configurations were used to evaluate the effects of a nonrandom distribution in habitat types and differences in patch size on the outcome of DA and CA analyses. In landscape 1, habitat patches were of equal size (10 ha) and randomly distributed in the study area. In landscape 2, habitat patches were also of equal size (10 ha), but the spatial distribution of habitat types was not random, with habitat H being juxtaposed to habitat A 95% of the time. To create this configuration, patches of habitat A were randomly distributed across the landscape and then adjacent patches were randomly assigned to habitat H so that 95% of habitat H was juxtaposed to habitat A. Remaining patches of habitat H and the six other habitat types were then randomly distributed across the rest of the landscape. To increase the spatial association between habitats A and H, landscape 3 was created by making habitat A entirely surrounded by habitat H. In this simulated landscape, both habitats A and H occurred in patches half the size (5 ha) of other habitat patches (10 ha). Landscapes 2 and 3 thus allowed us to assess the influence of a spatial association between the most- (A) and least-preferred (H) habitat types (see below). Landscapes 4 and 5 were similar to landscapes 2 and 3, but habitat A occurred in association with habitat C (a habitat that was used directly in proportion to availability), not H. In landscape 6, habitat types were randomly distributed across the study area, but the patch size of habitat A was four times (40 ha) the size of other habitat patches (10 ha).

Movements of 30 individuals were simulated in each of the six landscapes based on 40-km² ellipsoidal home ranges that were randomly positioned in the study area. Within each home range, we distributed 1000 animal locations among the eight habitat types in the following manner: habitats A (24.0%) and B (18.3%) were used

in greater proportion than availability; habitats C, D, E, and F were used in equal proportion to availability (12.5%); and habitats G (6.7%) and H (1.0%) were used in lower proportion than availability. Animal locations were randomly positioned within habitat patches, regardless of distance from edge. In other words, we intentionally simulated animal observations that violated the null hypothesis of no habitat selection to see how the DA and CA results would compare.

Empirical data

Habitat selection was assessed for 17 adult female moose (*Alces alces*) tracked with GPS telemetry collars during the summers of 1996 ($n = 6$), 1997 ($n = 7$) and 1998 ($n = 4$) in the boreal forest of Quebec, Canada (47°15' N, 71°20' W). Detailed information about the study area can be found in Dussault et al. (2001b). Capture methods were described in Dussault et al. (2004). Collars were programmed to record a location every 4 h, and GPS locations were differentially corrected to provide a horizontal position error of <15 m and <35 m, 50% and 95% of the time, respectively (Dussault et al. 2001a).

Habitat patches were classified into six categories: mixed and deciduous stands with shade-intolerant hardwoods <20 yr old (MI10); mixed and deciduous stands with shade-intolerant hardwoods 20–40 yr old (MI30); mixed and deciduous stands with shade-intolerant hardwoods >40 yr old (MI50); deciduous and mixed stands with tolerant hardwoods >40 yr old (DT50); coniferous stands >20 yr old (C30); and non-regenerated or non-forested areas (OTHER). Vegetation descriptions were based on 1:20 000 forest maps published by the Quebec Ministry of Natural Resources in 1992. These maps had been elaborated from 1:15 000 aerial photographs taken in 1990. Considering mean patch sizes of the habitat types (Table 1), the accuracy of moose positions provided by the GPS collars was adequate to measure habitat selection in that landscape. Map reliability was verified by comparing map classifications with field

measurements in summer 1997 (Dussault et al. 2001b). In order to facilitate interpretation of results, the extent of edge shared between all pairs of habitat types was used as a measure of the spatial association between habitat types (McGarigal and Marks 1994). Edge density was calculated for each habitat type as the number of meters per unit area bordering another habitat type.

Data analysis

For CA, log ratios were calculated for each individual using habitat H (simulations) or habitat C30 (empirical data) as the denominator in the ratios. In the case of an available habitat type that was not used by an individual, values of 0% use were replaced by 0.001% to allow calculation of log ratios. Habitat availability was calculated for each individual as the proportion of habitat types within the minimum convex polygon enclosing all observations (Mohr 1947). The series of log-ratio differences consisted of a vector with $n - 1$ components (where n = number of habitat types), and the mean vector for all animals was compared to a zero vector using a MANOVA. A matrix of pairwise comparisons between habitat types was used to establish the rank order of preference, with a rank of one indicating the highest preference. Habitat types between which log ratios did not differ received the mean rank. To determine which habitat types were selected or avoided, proportion use was compared to proportion availability with paired t tests on $\ln(x)$ -transformed data.

For DA, mean distance of locations to the nearest patch of each habitat type was calculated for each individual. This vector of mean observed distances (\mathbf{u}_i) was compared to the mean distances obtained using locations randomly distributed within the home range of each individual (\mathbf{r}_i). Random locations provided mean distances to each habitat type under a null hypothesis of no selection. A vector of ratios (\mathbf{d}_i) was created for each individual by dividing \mathbf{u}_i by \mathbf{r}_i . A MANOVA using $\ln(x)$ -transformed ratios as dependent variables was used to determine whether mean vector \mathbf{d}_i differed from 0 (no selection). Year was included as a random factor in the analysis with empirical data. As with CA, habitat types were ranked in order of preference by making pairwise comparisons of $\ln(x)$ -transformed ratios with t tests. To determine which habitat types were selected or avoided, paired t tests were used to compare $\ln(x)$ -transformed observed distances with random distances by habitat type. Similar numbers of random locations were used as the numbers of animal locations (i.e., 1000 locations in simulations and 232–609 locations with moose data). Random locations were generated using Hawth's Analysis Tools 2.10 (Beyer 2004). Distances between animal or random locations and habitat patches were calculated using

ArcMap 9.0 (ESRI 2000). All statistical tests were performed with the SAS program version 8.2 (SAS Institute 1989) with a significant probability level set at 0.05.

Simulations results

Habitat selection was detected using both CA and DA in the six simulated landscapes (CA, $F_{7,23} > 1446.9$, $P < 0.0001$; DA, $F_{8,22} > 31.8$, $P < 0.0001$). Results obtained with CA were the same no matter which landscape was considered (Fig. 1), with habitat types A and B selected, habitat types G and H avoided, and other habitat types used proportionately to availability. In addition, CA always provided the same rank order of preference ($A > B > C = D = E = F > G > H$). Results obtained with the DA method varied according to the simulated landscape, with both the spatial arrangement and relative size of habitat patches influencing the outcome of habitat selection analyses. The DA method only provided results similar to CA for landscape 1 where habitat patches of equal size were randomly distributed across the landscape (Fig. 1).

The greatest discrepancies between results of CA and DA were obtained using landscapes 3 and 5 where the most-preferred habitat type (A) was entirely surrounded by another habitat type (H or C; Fig. 1). In such a situation, both the habitat ranking and the determination of selection or avoidance of habitat types were affected. Proximity of the highly preferred habitat type (A) to the least-preferred habitat type (H), respectively, led to reduced and increased distances between animal locations and habitat type H and habitat type A compared to distances obtained with a random distribution of habitat types. In landscape 2, where 95% of the patches of the least-preferred habitat type (H) were juxtaposed to the most-preferred habitat type (A), the DA method attributed a higher ranking to habitat H than habitat G (Fig. 1). When habitat H entirely surrounded habitat A, this resulted in inability to discriminate between habitat types A and B in terms of ranking and in the conclusion that although habitat G was avoided, habitat H was used in proportion to availability (Fig. 1). Compared to landscapes 2 and 3, effects of a nonrandom distribution of habitat types were even greater in landscapes 4 and 5 where habitat type C (used in proportion to availability) was spatially associated with habitat type A. The juxtaposition of habitat C to habitat A in 95% of the cases and creating habitat A entirely surrounded by habitat C provided the following results; habitat type C was incorrectly found to be positively selected and to be ranked second, above habitat type B (Fig. 1).

Whereas varying patch size of the preferred habitat type did not modify the rank order of habitat types, patch size influenced determination of selection or

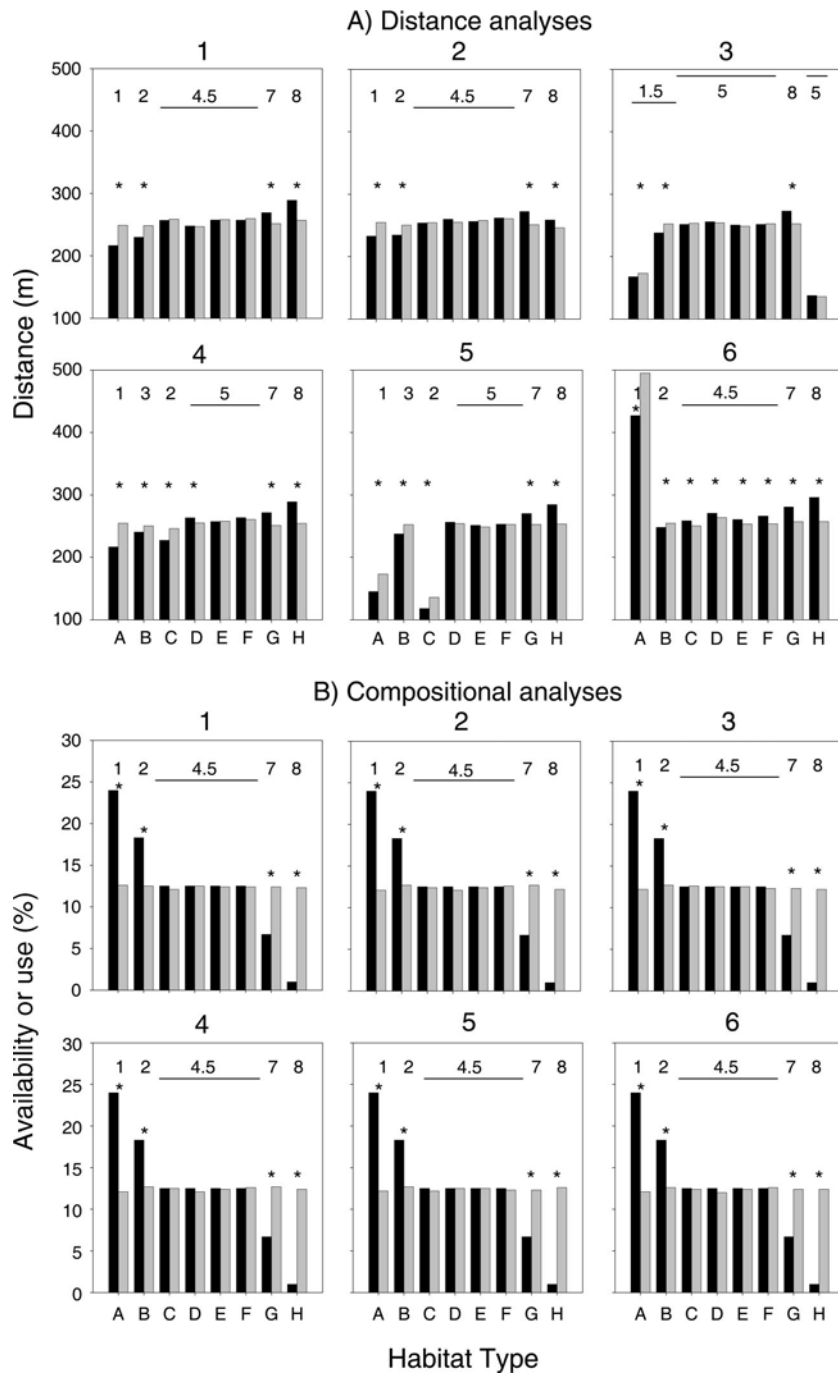


FIG. 1. Results of (A) distance analyses (DA) and (B) compositional analyses (CA) conducted on a series of landscapes (1–6) using various habitat configurations. Values are observed (black bars) and random (gray bars) distances for DA and percentage use (black) and percentage availability (gray) for CA. Asterisks indicate a significant difference ($P < 0.05$) between observed and random distances (DA) or use and availability (CA). Habitat rankings are indicated in each graph above histograms. Landscape 1: habitat types are randomly distributed, and habitat patches have equal size. Landscape 2: habitat H is juxtaposed to habitat A >95% of the time, and habitat patches have equal size. Landscape 3: habitat A is entirely surrounded by habitat H, and patches of both habitats A and H are half the size of other habitat patches. Landscape 4: habitat C is juxtaposed to habitat A >95% of the time, and habitat patches are of equal size. Landscape 5: habitat A is entirely surrounded by habitat C, and patches of both habitats A and C are half the size of other habitat patches. Landscape 6: habitat types are randomly distributed, and patches of habitat A are four times the size of other habitat patches.

avoidance of habitat types (Fig. 1). In landscape 6 where the preferred habitat type patches were four times the size of other habitat patches, observed distances to all other habitat types were higher than those obtained when all habitat types were of equal size, and we concluded that animals avoided habitat types that were actually used in proportion to availability (C to F; Fig. 1).

Empirical data results

CA and DA analyses both indicated that moose in the Jacques-Cartier Park did not use habitat types randomly (CA, $F_{5,8} = 43.6$, $P < 0.0001$; DA, $F_{6,22} = 23.7$, $P < 0.0001$). Results of DA and CA differed both in terms of use/availability comparisons by habitat type and rank order of preference. CA only indicated avoidance of habitat type C30 by moose whereas DA only indicated positive selection of habitat type MI10 by moose. Habitat ranking also differed between the two methods:

$$\text{CA: } \overbrace{\text{MI10}=\text{MI50}}^{1.5} > \overbrace{\text{MI30}=\text{DT50}=\text{OTHER}=\text{C30}}^{4.5}$$

$$\text{DA: } \overbrace{\text{MI10}}^1 > \overbrace{\text{MI30}}^2 > \overbrace{\text{MI50}=\text{DT50}=\text{OTHER}=\text{C30}}^{4.5}$$

In particular, habitat type MI30 was ranked higher using DA (2.0) than with CA (4.5) while habitat type MI50 was ranked lower with DA than CA (4.5 vs. 1.5).

Habitat types had different patch sizes and were not randomly distributed relative to each other within the study area, as revealed by examination of contrasted edge density between pairs of habitat types (Table 1). All habitat types except DT50 shared >50% of their edge density with the most abundant habitat type, i.e., C30. The habitat types for which CA and DA provided different rankings were mostly associated with MI50 and MI10 (MI30) and with MI10 and OTHER (MI50), in addition to C30.

Conclusions

Our simulations aimed to evaluate the correspondence and ease of interpretation of results obtained with DA to those obtained with a classification-based approach using the same data sets. DA results were evaluated using CA with the corresponding null hypothesis as a standard for comparison. With this in mind, the statistical procedures used in DA and CA are equivalent with the only noticeable difference between the two methods being the nature of the dependent variables (log ratios vs. Euclidean distances). As a result, the two methods can be compared directly (Conner et al. 2003). Because we neither simulated location error nor edge preference, we expected DA to provide results

similar to CA across simulated landscapes. Our simulations clearly indicated that the DA method can provide unexpected results when a preferred habitat type is spatially associated with another habitat type or when there is a systematic difference in patch size between habitat types.

Contrary to CA, results obtained with DA varied across the simulated landscapes even when proportionate use and availability of habitat types did not change. DA provided results similar to CA in only one of the six simulated landscapes. This indicates that interpretation of DA results is not straightforward and that it is not possible to relate the habitat ranking obtained with DA to resource preference *sensu* Aebischer et al. (1993) unless habitat types are randomly distributed and occur in similar sizes and shapes. Such conditions, however, are unlikely in natural environments. Species distributions are inherently nonrandom across landscapes since distribution depends on variables acting at the large scale, such as climate, or at smaller scales, such as topography, slope exposition, hydrology, nitrogen availability, and other factors (Senft et al. 1987, Pastor et al. 1998, Pastor et al. 1999). As an example, the majority of MI10 stands in our moose study area (Jacques-Cartier Park) originated from a spruce budworm (*Choristoneura fumiferana*) outbreak. This type of disturbance affects relatively old coniferous trees and thus is most likely to occur within mature coniferous (C30) or mixed stands (MI50), which resulted in >90% of the MI10 edge being shared with these two habitat types. Also, deciduous stand types are often found at lower elevations than coniferous stands, which explains why the degree of association between C30 and DT50 was <1%.

As demonstrated by the simulations, the spatial relationship between two habitat types may, under certain circumstances, create two effects with DA. When a highly preferred habitat type is associated with an avoided habitat type, mean distances to the preferred and avoided habitat types are, respectively, higher and lower than expected based on a random distribution of habitat types. With a very high degree of spatial association, the effect was great enough to erroneously conclude that the avoided habitat type was used in proportion to availability.

Differences in patch size between habitat types also influenced the assessment of habitat selection in DA simulations. When a highly preferred habitat type occurred in larger patch sizes than other habitat patches, distance to other habitat types was consequently high compared to that obtained using habitat patches of similar size (landscape 1). Indeed, if animals are not attracted to edges, relatively large habitat patches will necessarily result in high distances between animal locations and surrounding habitat types. As a result, mean

distances of locations to other habitat types should be overestimated compared to random locations because the mean is calculated based on a large proportion of locations in that habitat type. The opposite effect would be expected if an avoided habitat type occurs in larger patches than other habitat types. We also expect similar effects if a habitat type occurring in small patches relative to other habitat types is highly preferred or avoided. Again, as demonstrated in our moose study area, habitat types are not expected to have similar sizes. Even if habitat maps are often constructed from grid cells of the same size (e.g., satellite images), cells of the same habitat type that are juxtaposed are considered as a single habitat patch. Of course, one could expect differences in patch shape to have an effect similar to differences in patch size, with a preferred habitat type composed of long, irregular patches providing reduced distances to surrounding habitat types and the converse situation occurring with round preferred patches.

Using the empirical moose data, both CA and DA indicated that habitat selection occurred; however habitat ranking and conclusions regarding direction of habitat selection differed for the two methods. Habitat type MI30 was ranked 2 with DA and 4.5 with CA. We find this result surprising because Dussault et al. (*in press*) found MI50 to be significantly preferred over MI30 at the home-range scale during four time periods, with MI30 always occupying the lowest rank in preference. Furthermore, neither MI50 nor MI30 were found to be preferred at the landscape scale. We speculate that MI30 benefited from spatial association with MI50 and MI10, the two habitat types that were most preferred by moose, which resulted in a reduced mean distance to that habitat type for animal locations compared to random locations in DA. Differences in habitat patch sizes between habitat types may also have contributed to this discrepancy. The fact that habitat C30 was found to be avoided using CA but not DA could be related to its close association with MI10 and MI50. This association was not high enough however to modify the preference ranking, possibly because C30 was the most available habitat type in the study area (37.6%). Habitat MI10 was identified to be the most preferred habitat type using DA whereas CA results indicated that MI10 was not preferred over MI50. We speculate that MI10 obtained a higher rank than MI50 in DA because moose were found in edges bordering on habitat MI10 when using MI50 or C30 stands (Dussault et al., *in press*). Although difficult to demonstrate, spatial arrangement and differences in patch sizes between habitat types likely explain part of the differences between CA and DA in our empirical data. Nevertheless, these discrepancies invite caution in interpretation of DA results.

Our results should not be considered a plea for the exclusive use of CA. There exist several other classi-

fication-based approaches to study resource selection (Manly et al. 1993, Arthur et al. 1996, Cooper and Millsaugh 1999, Boyce et al. 2002). One major advantage of both DA and CA over other analytical approaches is the ability to assess habitat selection at various spatial scales. Our conclusions, however, only apply to third-order habitat selection as described by Johnson (1980). At the landscape scale (i.e., second-order selection), for example, spatial association of an avoided and a preferred habitat type could lead to the conclusion that the poor habitat type was selected using CA. Also, we did not consider any location error in our simulations and, as demonstrated by Conner et al. (2003), results obtained with CA may have been different using a high location error relative to patch size. When GPS telemetry is utilized, however, locations are accurate (approximately <20 m, 95% of the time), and misclassification is unlikely to affect the results if sample size is adequate.

We conclude that results obtained with the DA method cannot be easily linked to resource preference in natural environments. We recognize the potential influence of surrounding habitat types on habitat selection but also point out that results of the DA method, as proposed by Conner et al. (2003), should be interpreted cautiously in natural landscapes where habitat types are not spatially independent, regardless of decisions made by animals. The DA method could be improved by correcting observed distances for variation in habitat patch size and spatial correlation between habitat types by assessing these parameters using random locations. Until an improved method is proposed, we suggest restricting use of the DA method to specific questions, such as detecting edge preference within habitat types, i.e., by comparing distances of animal vs. random locations within a particular habitat type. Furthermore, DA could be useful to determine whether animals are closer than expected by chance to any particular feature in the habitat. However, it should be evident that being located closer to a habitat feature than expected does not automatically indicate that this feature is preferred.

Acknowledgments

We are especially grateful to M. Conner for his constructive and relevant comments. We also thank D. Fortin and an anonymous reviewer for helpful criticism on the manuscript, L. Breton for assistance in field work, G. Daigle from the Service de Consultations Statistiques of Laval University for assistance with statistical analyses, and W. J. King for carefully reviewing the manuscript.

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.

- Arthur, S. M., B. E. Manly, L. L. McDonald, and G. W. Garner. 1996. Assessing habitat selection when availability changes. *Ecology* **77**:215–227.
- Beyer, H. L. 2004. Hawth's analysis tools for ArcGIS, version 2.10. (<http://www.spatial-ecology.com/htools>)
- 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.
- Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. *Ecological Modelling* **157**:281–300.
- Conner, L. M., M. D. Smith, and L. W. Burger. 2003. A comparison of distance-based and classification-based analyses of habitat use. *Ecology* **84**:526–531.
- Cooper, A. B., and J. J. Millspaugh. 1999. The application of discrete choice models to wildlife resource selection studies. *Ecology* **80**:566–575.
- Dussault, C., R. Courtois, J. Huot, and J.-P. Ouellet. 2001b. The use of forest maps for the description of wildlife habitats: limits and recommendations. *Canadian Journal of Forest Research* **31**:1227–1234.
- Dussault, C., R. Courtois, J.-P. Ouellet, and J. Huot. 2001a. Influence of satellite geometry and differential correction on GPS location accuracy. *Wildlife Society Bulletin* **29**:171–179.
- Dussault, C., J.-P. Ouellet, R. Courtois, J. Huot, L. Breton, and H. Jolicoeur. *In press*. Linking moose habitat selection to limiting factors. *Ecography*.
- Dussault, C., J.-P. Ouellet, R. Courtois, J. Huot, L. Breton, and J. Larochelle. 2004. Behavioural responses of moose to thermal conditions in the boreal forest. *Ecoscience* **11**:321–328.
- ESRI. 2000. Using ArcMap. Environmental Systems Research Institute, Redlands, California, USA.
- Garshelis, D. L. 2000. Delusions in habitat evaluation. Pages 111–164 *in* L. Boitani and T. K. Fuller, editors. *Research techniques in animal ecology*. Columbia University Press, New York, New York, USA.
- Helzer, C. J., and D. E. Jelinski. 1999. The relative importance of patch area and perimeter–area ratio to grassland breeding birds. *Ecological Applications* **9**:1488–1458.
- Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. *Ecology* **61**:65–71.
- Manly, B. F. J., L. L. McDonald, and D. L. Thomas. 1993. *Resource selection by animals: statistical design and analysis for field studies*. Chapman and Hall, New York, New York, USA.
- McGarigal, K., and B. J. Marks. 1994. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Reference manual. Forest Science Department, Oregon State University, Oregon, USA.
- Mohr, C. O. 1947. Table of equivalent populations of North American small mammals. *American Midland Naturalist* **37**:223–249.
- Pastor, J., Y. Cohen, and R. Moen. 1999. Generation of spatial patterns in boreal forest landscapes. *Ecosystems* **2**:439–450.
- Pastor, J., B. Dewey, R. Moen, D. J. Mladenoff, M. White, and Y. Cohen. 1998. Spatial patterns in the moose–forest–soil ecosystem in Isle Royale, Michigan, USA. *Ecological Applications* **8**:411–424.
- SAS Institute. 1989. *SAS/STAT user's guide*. SAS Institute, Cary, North Carolina, USA.
- Senft, R. L., M. B. Coughenour, D. W. Bailey, L. R. Rittenhouse, O. E. Sala, and D. M. Swift. 1987. Large herbivore foraging and ecological hierarchies. *BioScience* **37**:789–799.
- Ecology*, 86(11), 2005, pp. 3125–3129
© 2005 by the Ecological Society of America

A COMPARISON OF DISTANCE-BASED AND CLASSIFICATION-BASED ANALYSES OF HABITAT USE: REPLY

L. Mike Conner,¹ Mark D. Smith,² and L. Wes Burger, Jr.²

Dussault et al. (2005) suggest that a Euclidean distance approach (DA) to the analysis of habitat use is difficult to interpret, and they provide simulation analyses to illustrate that variation in patch size and spatial correlations among patch types may influence habitat rankings when using a DA. Herein, we clarify the difference between DA and classification-based approaches (CA) and their interpretation. We also provide evidence that the effects of landscape configuration may result in similar interpretive issues with CA and suggest that the effects of landscape configuration should be interpreted using knowledge of the species under study. Additionally, we suggest that violating a classification-based null hypothesis for the purpose of comparing habitat rankings between DA and CA does not provide an appropriate comparison. Finally, we offer a slightly different interpretation of the moose (*Alces alces*) habitat selection data presented by Dussault et al. (2005).

The null hypothesis associated with a DA is that the distances between animal locations and the nearest patch of each habitat are equal to those of random locations. If the difference between these distances is sufficiently large relative to the variance, then the null hypothesis is rejected (Conner and Plowman 2001, Conner et al. 2003). In Conner et al. (2003), we suggested that animal locations that fall closer to a habitat feature than expected are selected, whereas locations that are farther from habitats than expected are avoided. It seems that defining habitat as either selected or avoided may be a source of confusion. A more strict interpretation of DA results is likely a less ambiguous approach (i.e., restricting inference to animals being closer to or farther away from habitat features than

Manuscript received 13 May 2005; accepted 17 May 2005.
Corresponding Editor: A. M. Ellison.

¹ Joseph W. Jones Ecological Research Center, Route 2, Box 2324, Newton, Georgia 39870 USA.
E-mail: mconner@jonesctr.org

² Department of Wildlife and Fisheries, Box 9690, Mississippi State University, Mississippi State, Mississippi 39762 USA.

expected). These statements will be true regardless of spatial configuration of patches or patch sizes; locations will be closer to or farther away from a habitat than expected or they will be randomly located relative to the habitat. There should be little room for misinterpretation of the results. Why animals are closer to some habitats and farther away from others should be explained based on the researcher's understanding of the organism under study. Further, this interpretation of a DA is analogous to accepted methods of interpreting a CA (i.e., a habitat is used more or less than expected and it is left to the researcher's knowledge of the study organism to explain the results).

Dussault et al. (2005) provide an example using simulated data to illustrate that DA can rank a habitat differently due to spatial association with other habitats. However, the same spatial correlation can affect conclusions when using a CA. Consider the examples put forth in Dussault et al. (2005) where an avoided habitat (H) was strongly associated with a preferred habitat (A); these examples were referred to as landscapes 2 and 3. Dussault et al. (2005) used these landscapes to show that a spatial correlation can affect habitat ranking when using a DA. Similar to Dussault et al. (2005), we provide a simple example to demonstrate that spatial correlation of habitats may also affect results obtained from using a CA. We do not intend for our example to indicate that classification-based analyses are flawed, but rather to illustrate that any habitat analysis can be affected by spatial correlation, depending upon scale of investigation. We also present an empirical example (Conner and Leopold 1996) where relative habitat rankings using a CA were affected by a spatial correlation between two habitats.

Consider a simple landscape in which three habitat types occur (Fig. 1). Assume that a preferred habitat (A) completely surrounds an unusable habitat (B) and that A is completely surrounded by an unusable habitat (C). For simplicity sake, further assume that each block of habitat A is entirely used by a single animal, and that no animals are ever located in habitats B or C. Habitat B represents only 4% of the landscape, but because it is completely surrounded by a preferred habitat, its composition within home ranges is 11.1%. Thus, habitat B is used greater than available at Johnson's (1980) second order of selection, despite the fact that animals never used the habitat. Habitat B would be avoided at Johnson's (1980) third order of selection, while habitat C would be avoided at both orders of selection. Whereas this example may be overly simplistic, it clearly illustrates that spatial correlation of habitats may affect results of CA as well as DA.

A similar scenario was observed by Conner and Leopold (1996) when using a CA to analyze bobcat (*Lynx rufus*) habitat use data obtained via radiotelemetry. In

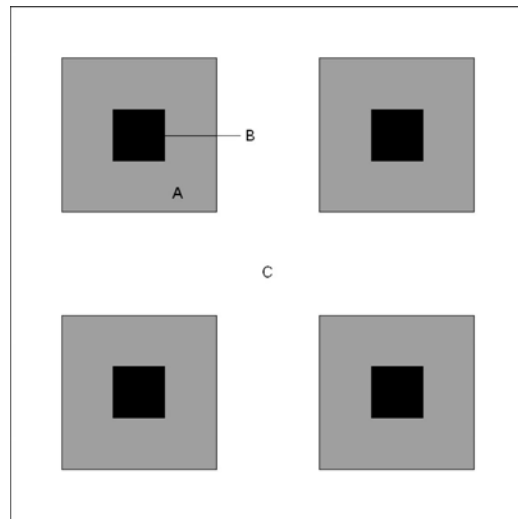


FIG. 1. Simulated landscape containing three habitats. Availabilities of habitats are $A = 32\%$, $B = 4\%$, and $C = 54\%$. If each area of A is completely used by an individual animal and animals do not use habitats B and C, the percentage of habitats within animal home ranges will be: $A = 89\%$, $B = 11\%$, and $C = 0\%$. Because B is associated with A, B appears to be selected more than available at Johnson's (1980) second order of selection (i.e., habitat composition of the home range relative to study area availability).

their study, an analysis of habitat selection at Johnson's (1980) second order of selection indicated that percentage of the home range composed of mature pine (*Pinus* spp.) exceeded availability of mature pine on the study area. In contrast, Johnson's (1980) third order of analysis indicated that mature pine stands were avoided. Conner and Leopold (1996) concluded that because a selected habitat (young pine plantations) was generally located within mature pine, that mature pine was preferred at Johnson's (1980) second order of selection.

When habitat rankings differ at different spatial scales, researchers should consider spatial correlation among habitats as one of several potential causes for this discrepancy. For example, if a given habitat is selected at Johnson's (1980) third order of selection but is avoided at Johnson's (1980) second order of selection, a spatial correlation among patches may be responsible. A similar discrepancy was evident in Conner and Leopold (1996) when using a CA in an analysis of bobcat habitat. If spatial correlation or configuration of habitats is thought to be affecting a habitat analysis, knowledge of the species in question should be relied upon to interpret the data further. Conner and Leopold (1996) had data to indicate that prey abundance was low in mature pine stands and data to reveal the strong association between mature pine stands and a preferred habitat (young pine plantations). Thus, they concluded

that selection for mature pine at Johnson's (1980) second order was likely due to a spatial association between mature pine stands and the young pine plantations inherently embedded within these mature stands. In the rare case when no prior knowledge is available, results should be interpreted with caution and all competing explanations should be discussed and suggested as hypotheses for future work.

Dussault et al. (2005) state that they used the same approach as Conner et al. (2003). We agree regarding the example provided using the animal-derived moose data. However, we suggest this is not the case with the simulated data. In the moose example, real animal locations were used, and researchers were free to define selection or avoidance to fit the analysis approach. Thus, selection could be based on percentage use relative to availability when using a CA or based on distances of moose and random (i.e., expected) locations to habitats when using a DA. Study animals performed the actual selection of habitats, and it was left up to the scientist to render an interpretation. This was not the case with simulated data.

Dussault et al. (2005) used simulations to compare relative rankings among eight habitats between CA and DA. In Dussault et al. (2005), locations were assigned to individual habitats under varying landscapes configurations such that the null hypothesis of no habitat selection was violated. However, the violation of the null hypothesis was simulated under a CA, not a DA. This is problematic because the null hypotheses differ between CA and DA.

The null associated with CA is that the percentage use of habitats is equal to their percentage availability. Here, use and availability are determined by classification. Because locations were simulated to fall into certain habitats at a predetermined rate, the simulation was by definition, a simulation designed to assess a classification procedure. In other words, a classification-based simulation was used to assess a classification-based procedure with a classification-based null hypothesis. Because Dussault et al. (2005) used compositional analysis (Aebischer et al. 1993) as their CA and because compositional analysis is based on relatively well established statistical procedures (i.e., MANOVA), their CA performed as would be expected (i.e., rankings of habitats were consistent with the way in which the null hypotheses were violated).

In contrast, the null associated with DA is that distances between animal locations and habitat types are equal to expected (i.e., random) distances. This difference in null hypotheses is not trivial. A simulation based on CA does not properly evaluate a distance-based statistic because we can not ascertain what the mean distance between animal locations and a given habitat *should be* under the false null associated with

CA (i.e., when percentage use \neq percentage availability). For example, if a habitat is simulated as 25% available and 50% used, we can say the use : availability ratio is expected to be 2:1; however, the use distance: random distance ratio that would be expected using this simulation is unknown. Likewise, a violation of the null under DA would not yield a predictable use : availability ratio for evaluating CA. Thus, the simulations by Dussault et al. (2005) do not provide an appropriate standard to evaluate DA. Because the nulls differ between DA and CA methodologies, we suggest that DA should be evaluated by violating the null hypothesis that distances between animal locations and habitats are equal to expected distances obtained from random locations. It may be possible, however, to directly compare CA and DA by simulating animal movement rates and turning angles (Turchin 1998).

We realize that Bingham and Brennan (2004) similarly compared CA to DA. However, because their simulations were generated under a true null hypothesis, CA and DA were expected to provide the same results if Type I error rates were the same. In other words, randomly generated locations (the numerator in the distance analysis representing distance from animal to habitat) should equal the random distance to habitats (the denominator representing the null model in a DA analysis). To simplify, a random distance divided by a random distance will, on average, equal one. Thus, because simulated locations were generated under a true null, their simulations were acceptable for both CA and DA analyses.

The simulations used by Dussault et al. (2005) assume that an animal's use of a given habitat is independent of surrounding habitat patches. By ensuring that animals select habitat independently of surrounding habitats, Dussault et al. (2005) assumed that the quality of a patch was unaffected by surrounding patches. In contrast, Conner et al. (2003) state "Because a distance-based approach uses distances to the nearest habitat feature of interest regardless of whether the feature is within the home range or study area boundary, surrounding habitats have the opportunity to influence the results of the analysis." Thus, there is no assumption that animal locations fall within a given patch independently of surrounding patches. We consider this a benefit of using distances in analyses of habitat selection.

We suggest that assuming usage of a given habitat patch is independent of surrounding patches is likely unrealistic. Johnson (1980) emphasized that habitat selection is a hierarchical process, with selection at one level of this hierarchy based on prior selection at a more general level. If so, it is reasonable to assume that use of a given patch is influenced by surrounding patches. For example, does the presence of a poor hab-

itat surrounding a good habitat reduce the usefulness of the good habitat? Most would agree that it does. Indeed this concept is often used to describe absence of a species in what appears to be suitable habitat. For example, Cram et al. (2002) reported that suitability of a given patch in combination with amount of other suitable habitat in the surrounding landscape influenced occupancy and relative abundance of northern bobwhite (*Colinus virginianus*) in Arkansas. Cram et al. (2002) concluded that "Small isolated treatment stands in the ONF may offer suitable habitat structure. . . , but due to lack of suitable surrounding habitat, these stands are too small to support viable bobwhite populations."

Whereas Dussault et al. (2005) analyzed moose data to provide a non-simulated, animal-based example of using a CA and DA, our interpretation of their results differs slightly from that presented. Dussault et al. (2005) state that MI50 was selected over MI30 at Johnson's (1980) third order of selection (i.e., home range scale) using CA and that MI30 was ranked as 2 using DA but 4.5 using CA (rank = 1 indicates most selected; rank = 5 indicates least selected). They further state that neither MI30 nor MI50 were found to be selected at Johnson's (1980) second order of selection (i.e., landscape scale) when using a CA. We suggest that a spatial correlation of MI30 and MI50 with less selected habitats (e.g., C30) may be responsible for the avoidance at the landscape scale. This is particularly true given that MI50 shared 51.3% of its edge with C30, the most avoided habitat at Johnson's (1980) third order of selection.

Dussault et al. (2005) speculated that the ranking of MI30 above MI50 (rank = 4.5) in the DA was because of the close association of MI30 with MI50 and MI10, the most selected habitats. Interestingly, the statement that MI50 and MI10 are most selected unquestionably assumes CA to be the correct approach. We suggest an alternative interpretation of the DA and CA comparison of moose data put forth in Dussault et al. (2005). We point out that MI30 actually shared more edge with C30 (54.1%), the lowest ranked (i.e., least selected) habitat using both CA and DA, than with MI10 and MI50 (combined shared edge = 32.2%). Further, whereas a spatial association of MI30 with a more selected habitat (e.g., MI50) may result in an improved ranking for MI30, a spatial association alone is not sufficient to elevate the rank of MI30 *above that of* MI50 (i.e., MI30 selected over MI50). In other words, a close association of an avoided habitat with a selected habitat can not result in an avoided habitat being ranked *above* a selected habitat based on that association alone when using a DA. We acknowledge that our experience with moose is limited, but given the results as presented, we must conclude that MI30 should be ranked above MI50 as indicated in the DA.

We do not advocate DA over CA; the tool to use should be based on objectives of the study and experience with the species. However, we remain convinced that the DA approach retains more information regarding effects of landscape physiognomy on animal locations than classification-based approaches. Every animal location occurs within a context that can be described with regard to patch type, patch characteristics (i.e., shape, size, and other factors), patch adjacencies, and landscape composition and configuration. Classification approaches use only a small part of this contextual information, patch type. Once a location is classified as belonging to a given habitat, remaining information associated with that location is lost (Conner et al. 2003). This does not occur with a DA. Further, DA permits use of habitat features that either lack area or are represented using models that lack area (e.g., while a road has area, it is typically represented thematically as a line that lacks area). Because more information is retained regarding the proximity of habitats to animal locations and because spatial correlation of habitats can affect both CA and DA analyses, it may be desirable to model the degree of spatial correlation of habitats and assess the effects of spatial correlations on analysis results. This would not be possible within the strict confines of a traditional CA approach but may be possible with DA. Using DA to model spatial correlation of habitats and the effects of this correlation on habitat selection should be a profitable area of research.

Acknowledgments

The Joseph W. Jones Ecological Research Center provided funding for this research. D. Miller, J. Cochrane, L. Smith, and G. Conner provided valuable comments that improved the quality of the manuscript.

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.
- Bingham, R. L., and L. A. Brennan. 2004. Comparison of type I error rates for statistical analyses of resource selections. *Journal of Wildlife Management* **68**:206–212.
- Conner, L. M., and B. D. Leopold. 1996. Bobcat habitat use at multiple spatial scales. *Proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies* **50**:622–631.
- Conner, L. M., and B. W. Plowman. 2001. Using Euclidean distances to assess nonrandom habitat use. Pages 275–290 in J. Millsbaugh and J. Marzluff, editors. *Radio telemetry and animal populations*. Academic Press, San Diego, California, USA.
- Conner, L. M., M. D. Smith, and L. W. Burger. 2003. A comparison of distance-based and classification-based analysis of habitat use. *Ecology* **84**:526–531.
- Cram, D. S., R. E. Masters, F. S. Guthery, D. M. Engle, and W. G. Montague. 2002. Northern bobwhite population and

- habitat response to pine–grassland restoration. *Journal of Wildlife Management* **66**:1031–1039.
- Dussault, C., J.-P. Ouellet, and R. Courtois. 2005. A comparison of distance-based and classification-based analyses of habitat use: comment. *Ecology* **86**:3119–3125.
- Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluating resource preference. *Ecology* **61**:765–771.
- Turchin, P. 1998. *Quantitative analysis of movement*. Sinauer Associates, Sunderland, Massachusetts, USA.
-