An evaluation of three statistical methods used to model resource selection

David M. Baasch [a,*], Andrew J. Tyre [a], Joshua J. Millspaugh [b], Scott E. Hygnstrom [a], Kurt C. Vercauteren [c]

[a] School of Natural Resources, University of Nebraska, Lincoln, NE 68583, USA
[b] Department of Fisheries and Wildlife Sciences, University of Missouri, Columbia, MO 65211, USA
[c] USDA/APHIS/WS/National Wildlife Research Center, Fort Collins, CO 80521, USA

ARTICLE INFO

ABSTRACT

The performance of statistical methods for modeling resource selection by animals is difficult to evaluate with field data because true selection patterns are unknown. Simulated data based on a known probability distribution, though, can be used to evaluate statistical methods. Models should estimate true selection patterns if they are to be useful in analyzing and interpreting field data. We used simulation techniques to evaluate the effectiveness of three statistical methods used in modeling resource selection. We generated 25 use locations per animal and included 10, 20, 40, or 80 animals in samples of use locations. To simulate species of different mobility, we generated use locations at four levels according to a known probability distribution across DeSoto National Wildlife Refuge (DNWR) in eastern Nebraska and western Iowa, USA. We either generated 5 random locations per use location or 10,000 random locations (total) within 4 predetermined areas around use locations to determine how the definition of availability and the number of random locations affected results. We analyzed simulated data using discrete choice, logistic-regression, and a maximum entropy method (Maxent). We used a simple linear regression of estimated and known probability distributions and area under receiver operating characteristic curves (AUC) to evaluate the performance of each method. Each statistical method was affected differently by number of animals and random locations used in analyses, level at which selection of resources occurred, and area considered available. Discrete-choice modeling resulted in precise and accurate estimates of the true probability distribution when the area in which use locations were generated was ≥ the area defined to be available. Logistic-regression models were unbiased and precise when the area in which use locations were generated and the area defined to be available were the same size; the fit of these models improved with increased numbers of random locations. Maxent resulted in unbiased and precise estimates of the known probability distribution when the area in which use locations were generated was small (home-range level) and the area defined to be available was large (study area). Based on AUC analyses, all models estimated the selection distribution better than random chance. Results from AUC analyses, however, often contradicted results of the linear regression method used to evaluate model performance. Discrete-choice modeling was best able to estimate the known selection distribution in our study area regardless of sample size or number of random locations used in the analyses, but we recommend further studies using simulated data over different landscapes and different resource metrics to confirm our results. Our study offers an approach and guidance for others interested in assessing the utility of techniques for modeling resource selection in their study area.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Resource selection is a valuable field of study in animal ecology. Conclusions drawn from resource selection studies have important implications because they often serve as guidelines for habitat management plans and habitat suitability indices (Garshelis, 2000).

Effective management and conservation of species requires an understanding of habitat requirements, well-guided techniques for collecting data, and robust methods for analyzing data. Animals select resources at different scales dependent upon what is actually available to them (Owen, 1972). What is perceived by the animal to be available is impacted by numerous factors that are too complex to measure fully, but include distance, barriers, inter- or intra-specific competition, risks, and habits or patterns of selection. While we may not be able to assess all the factors that an animal considers when using specific locations, the analyst's determination of areas of use and availability are important factors in determining how well a model actually represents the population

of interest, regardless of the statistical method used (McClean et al., 1998; Aebischer et al., 1993; Buskirk and Millspaugh, 2006).

Determination of what is considered available to the animal and the numbers of animals to include in the sample are recurring concerns that exist in analyses of use versus availability studies (Alldrege and Ratti, 1992; Leban et al., 2001; Buskirk and Millspaugh, 2006). In many studies, it is difficult to determine areas an animal did not use (Austin, 2002) and therefore resource selection analyses often involve a determination of available resources and the generation of random locations (Keating and Cherry, 2004; Johnson et al., 2006). In particular, radio-telemetry studies only document an animal's use patterns; this type of data is extremely common (Cooper and Millspaugh, 1999; Erickson et al., 2001; Manly et al., 2002; MacKenzie, 2006). Several definitions of availability exist in the literature, most of which are associated with the geographic range of a species, study area, home range, or localized areas around each use location (Johnson, 1980; Buskirk and Millspaugh, 2006). The researcher's choice of what is defined to be available to the animal can affect the results of the analysis (Johnson, 1980; McClean et al., 1998; Boyce et al., 2003). Some studies used the observed mobility of the study species (i.e., displacement distance between consecutive observations) to define available resources (e.g., Arthur et al., 1996; Cooper and Millspaugh, 1999), but until recently most quantitative methods did not allow for such flexibility and definitions of availability were more subjective. The number of animals observed in a study can also affect results of resource selection analyses. Alldrege and Ratti (1986) compared univariate and nonparametric approaches for modeling resource selection and reported methods used by Neu et al. (1974) and Quade (1979) performed well at sample sizes ≥20 animals with 50 locations/animal. Leban et al. (2001) reached similar conclusions when evaluating compositional analysis (Aebischer et al., 1993). Some studies have evaluated performance of statistical models at different levels of availability; however, the true underlying probability distribution of the selection patterns was not known (McClean et al., 1998; Phillips et al., 2004; Özesmi et al., 2006) or was assumed (Leban et al., 2001). We are unaware of any studies, however, that used simulated data generated according to a known probably distribution to examine the effects of sample size (number of animals), number of random locations, and areas of use and availability on results of multivariate techniques for modeling resource selection.

Several statistical techniques have been used to develop resource selection functions (RSFs). Through the 1990s, logistic regression was widely used for analyzing data for producing RSFs (Manly et al., 1993). Recent criticism of the use of logistic regression in use-availability studies, however, has raised substantial questions about the method (Keating and Cherry, 2004). In recent years, the use of discrete-choice models for generating RSFs has increased (Cooper and Millspaugh, 1999; McDonald et al., 2006; Thomas et al., 2006). One advantage of using discrete-choice modeling is that the model allows the researcher to develop a different choice set for each independent observation of use. Developing a separate set of random locations for each use location allows resources to change throughout the study period and helps ensure resources defined as available were accessible by the animal when selection of the use location occurred (Arthur et al., 1996; Cooper and Millspaugh, 1999).

In this study, we explore a new, maximum entropy approach for developing RSFs (Maxent, version 3.0.6, Phillips et al., 2005). Maxent is a machine learning method that has several aspects that make it well suited for modeling resource selection, and has been effective at making predictions or inferences from incomplete information in other domains, such as species distribution modeling (Phillips et al., 2004; Phillips and Dudík, 2008) and natural language processing (Berger et al., 1996). We are unaware, however, of any application of Maxent in studies of habitat or resource selection. Maxent estimates the RSF by finding the distribution of maximum entropy subject to the constraint that the expected value of each feature under this estimated distribution matches its empirical average (Phillips et al., 2006). The method of estimation used by Maxent is equivalent to finding the maximum likelihood distribution that is exponential in a linear combination of the features (i.e., Gibbs distribution, Phillips et al., 2004). The deterministic algorithms used in Maxent are guaranteed to converge to the optimal (maximum entropy) probability distribution and employ a regularization function to prevent algorithms from over-fitting the data (Phillips et al., 2006; Phillips and Dudík, 2008). Della Pietra et al. (1997), Collins et al. (2002), Dudík et al. (2004), Phillips et al. (2004, 2006), and Phillips and Dudík (2008) provide detailed information on Maxent, machine learning, and the underlying updating algorithms used in Maxent.

The most appropriate method for generating RSFs, assuming a single best method exists, can be determined using simulated data (Berger et al., 1999; Hirzel et al., 2001; Tyre et al., 2001). Our objectives were to: (1) compare the performance of discrete-choice, logistic-regression, and Maxent models in resource selection studies, (2) determine effects of sample size (number of animals) and number of random locations used in analyses for each modeling technique, and (3) determine effects of species' mobility and area defined to be available on performance of methods. If modeling methods cannot consistently estimate a known probability distribution based on simple theoretical models, application of the methods to real data is questionable at best, even if all statistical assumptions are met.

2. Methods

We used five steps to investigate the performance of three methods for estimating RSFs including: (1) generation of realistic environmental data; (2) generation of simulated species data (i.e., use locations) responding to direct environmental gradients according to a known probability distribution; (3) generation of random locations with respect to predefined choice sets; (4) selection of an appropriate statistical model to be evaluated; and (5) evaluation of methods.

2.1. Generation of realistic environmental data

DeSoto National Wildlife Refuge (DNWR), an irregularly shaped area of 3166 ha in Iowa and Nebraska, USA defined our study area. In our study area, the average size of terrestrial patches was 10.6 ha (range = 0.1–162.5 ha); <3% (7/265) of the terrestrial patches were >50 ha in size. We used ArcMap version 9.2 (ESRI, 2006) to convert the study area into a 30 m × 30 m raster that defined the 31,136 possible resource units within DNWR. We assumed the simulated species was not aquatic and did not use roads to increase fitness and removed these resource units from the study area. We classified all of the raster points by landcover type and calculated the distance of each point to road, water, and edge of wooded area. Distance to edge of woods was negative for all grid cells classified as wooded and positive for all other grid cells. Distance to road and distance to water were all ≥0.0.

Coefficients from an RSF model of data collected for white-tailed deer (*Odocoileus virginianus*) at DeSoto National Wildlife Refuge during 1991–1997 were used to generate the known selection distribution (Baasch, 2008). Each grid cell (i) had a known probability of use given by Eq. (1) (provided in Appendix A.2) where $\beta_1 = -0.3617$, $\beta_2 = 0.0708$, $\beta_3 = 0.0424$, $\beta_4 = 0.3814$, $\beta_5 = 0.2657$, $\beta_6 = -0.1932$, $\beta_7 = -0.3898$, $\beta_8 = -0.6384$, $\beta_9 = 0.2460$, $\beta_{10} = 0.06837$ and where x_{i1}–x_{i6} were six of the seven levels of the

categorical variable (landcover type). The final level of the land-cover variable (corn) was the reference category that was assigned a β-value of 0.0. Variables x_{i7}–x_{i10} were continuous distance measures (distance to road, water, edge of wooded area, and edge of wooded area squared).

2.2. Generation of use locations

We included 25 use locations per animal and varied sample size by including 10, 20, 40, or 80 animals in the sample (250, 500, 1000, and 2000 use locations, respectively) to test the sensitivity of each method to number of animals used in the analysis. Use locations were simulated observations of resource use generated according to the known underlying probability distribution described in Section 2.1. We generated use locations within 300 m, 600 m, or 1200 m of the previous use location to simulate species of lower mobility or species that have a home range smaller than the study area (e.g., turtle, deer; Fig. 1). Use locations were also generated across the study area to simulate species of higher mobility that could traverse the entire study area on a daily basis. We define "level of use" as the area in which use locations were generated (within 300 m, 600 m, or 1200 m of previous use location or across the study area).

For the first level of use (300 m), we generated use locations 1 at a time with replacement. The first use location for each animal was generated from the set of all possible resource units within the study area (Appendix A.1(1)). Use locations 2–25 for each animal were generated sequentially with replacement (Appendix A.1(2)). We generated these locations at a minimum distance of 90 m from the previous use location and within a 600 m × 600 m square area centered on the previous use location (i.e., the animal traveled between 90 m and 300 m prior to recording another use location). The probability of use for any subsequent resource unit was equal to the probability of the underlying raster, normalized within the predefined area. We used similar techniques to generate use locations for the second and third levels of use (600 m and 1200 m, respectively). For the study-area level of use, all use locations were generated simultaneously across the study area with replacement (Appendix A.1(1)). The probability of use for each resource unit was independent of previous use locations with this method of selection.

2.3. Generation of random locations

We generated random locations within areas defined to be available to the animal to represent resource units the animal could have chosen. We tested four levels of availability with respect to each use location (300 m, 600 m, 1200 m, and study area), to determine the best definition of availability for each statistical method tested. To allow direct comparisons, we tested the performance of all three statistical methods with the total number of random locations equal to either 5 × number of use locations, as suggested by McFadden (1978) for discrete-choice analyses, or 10,000 random locations total (default setting in Maxent). We generated 1250, 2500, 5000, or 10,000 random locations when the number of random locations = 5 × number of use locations, for sample sizes of 10, 20, 40, or 80 animals with 25 use locations per animal, respectively. We generated 40, 20, 10, or 5 random locations per use location for sample sizes of 10, 20, 40, or 80 animals, respectively when the total number of random locations = 10,000. For clarity, we only present the results from the 2 extremes of 10 and 80 animals; other results fell in between those 2 extremes. Random locations were chosen without replacement from the 30 m × 30 m raster-grid for each simulation trial (Appendix A.1(2)); the probability of selecting any resource unit was equal. Random locations were generated outside a 2-ha square region centered on the respective use loca-

tion, representing a "pseudo-error polygon" to account for error known to exist in telemetry studies, and within a square area centered on each use location with length and width equal to 2 times the level defined to be available to the animal. For example, if the level of availability was 300 m, we allowed the animal to choose from the set of resource units located >60 m from the use location and within a squared area ranging from 300 m to the north, south, east, or west of the use location (Fig. 1).

2.4. Statistical models

We stratified random locations with each use location chosen from the defined choice sets at each predetermined sample size for discrete-choice analyses (Appendix A.1(3), Eq. (1)). We also used a logistic-regression model (Eq. (2)) to analyze sets of data where the "choice set" was defined as all random locations located within a predetermined distance of any use location (Appendix A.1(4)). The model included the same variables as the model used to produce the underlying probability distribution (Baasch, 2008). In addition, we used Maxent to analyze sets of data and included the same terms as the other methods (Eq. (5)). We programmed Maxent to model linear terms only and entered "distance to edge of wooded area squared" by squaring these values within data sets prior to importing into Maxent.

For analyses involving discrete-choice models, we discarded all simulations that resulted in a coefficient for a landcover variable <−10.0, which indicated random locations were generated in a class of landcover in which no use locations were generated. Likewise, for Maxent we discarded all simulations in which an estimate for a coefficient was missing (null). By discarding these simulations, we avoided having averages for estimates of coefficients skewed by a few trials that resulted in an extreme estimate for a coefficient (e.g., −17.0) and were able to produce a complete set of coefficients for each simulation.

2.5. Performance measures

We performed 100 simulations for each statistical method at each level of use and availability and for each sample size resulting in 38,400 simulations of data. We compared the performance of the three statistical methods (logistic regression, discrete choice, and Maxent) by comparing the ability of each method to estimate the known probability distribution with a calibration curve and by using an analysis of area under the receiver operating curve (AUC). The use of AUC for comparing the predictive ability of models has increased in species distribution modeling and natural language processing (Hanley and McNeil, 1982; Bradley, 1997; Park et al., 2004; Fawcett, 2005), but its' application to studies of resource selection has only recently occurred (Boyce et al., 2002). The AUC is equivalent to the probability of a model ranking randomly chosen use locations higher than random locations, which is equivalent to the Wilcoxon test of ranks (Hanley and McNeil, 1982; Park et al., 2004; Fawcett, 2005). The AUC is the portioned area of correct classifications under the ROC curve so its value will always be between 0 and 1.0 and can be interpreted as a probability of correct classification or prediction. The ability to discriminate between use and random locations (i.e., performance) of various statistical methods can be directly measured and compared using AUC scores.

We compared the ability of the three statistical methods to estimate the original probability distribution using all use and random locations in each simulation. We used Eqs. (1) and (2) (Appendix A.2) to calculate estimated probabilities of resource selection across the study area for all 100 discrete-choice and logistic-regression models, respectively. We used linear regression with estimated selection as the predictor of known selection; each simulation produced one calibration curve with an intercept and slope. We

(A)

(B)

(C)

(D)
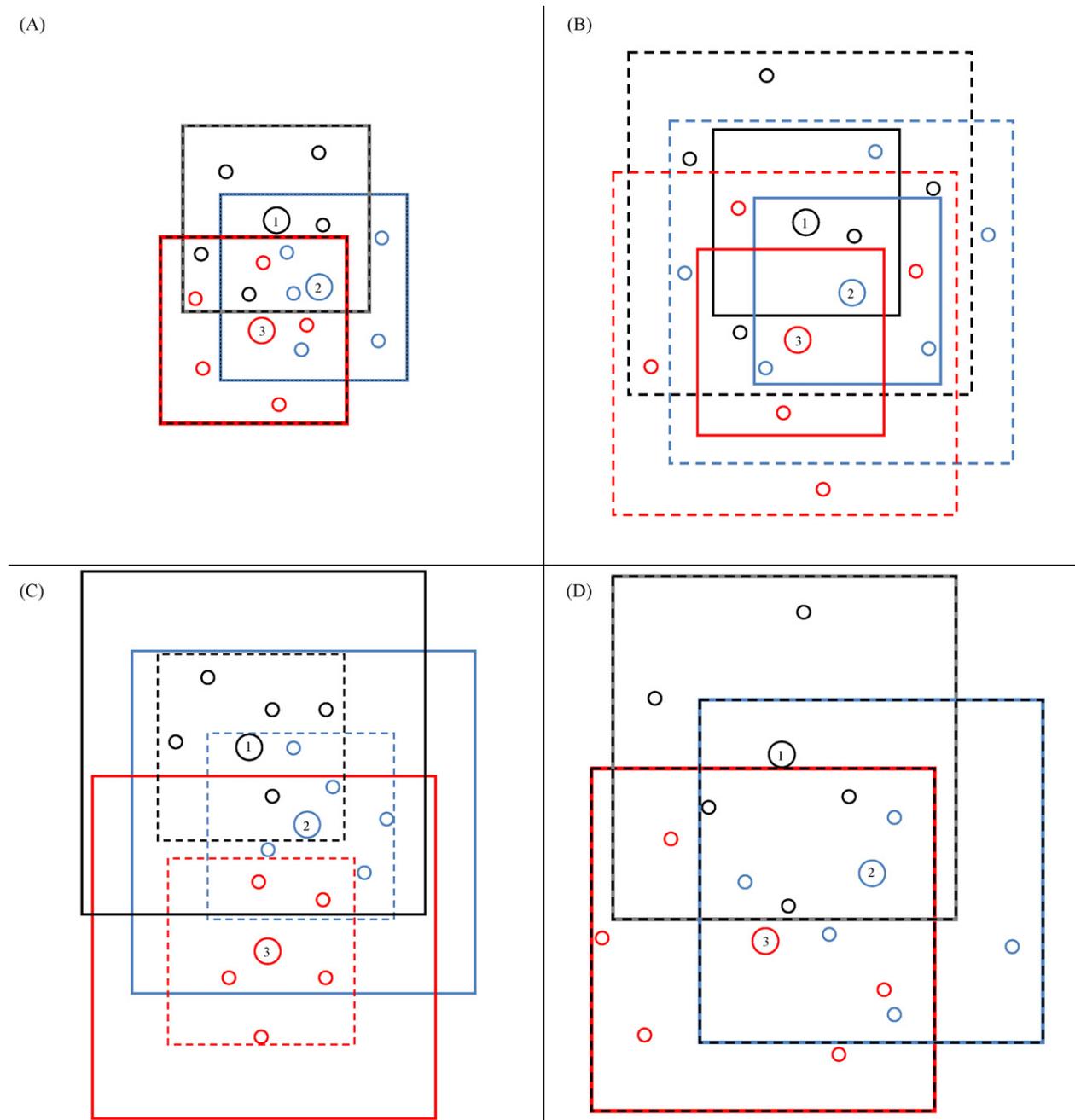
Fig. 1. An illustration of how we generated simulated locations of use (larger black, blue, and red numbered-circles) and random locations (smaller black, blue, and red circles) used in resource selection analyses. In all 4 examples (panels A, B, C, and D), we: (1) generated use locations 1 at a time and chose the first use location (1), according to a known probability distribution, from the set of all possible 30 m × 30 m resource units within the study area; (2) generated use locations 2 and 3 (of the 25/animal simulated in the study) within an area in which we determined use could occur (solid-line boxes), at a minimum distance of 90 m from the previous use location, and sequentially with replacement; and (3) generated 5 random locations/use location within the area we defined to be available (dashed-line boxes). (A) We generated use locations 2 (blue, numbered circle) and 3 (red, numbered circle) within a 600 m × 600 m square area centered on the previous use location (i.e., the animal traveled between 90 m and 300 m prior to recording another use location). We generated random locations within a 600 m × 600 m square area centered on each use location. (B) We generated use locations 2 and 3 within a 600 m × 600 m square area centered on the previous use location and generated random locations within a 1200 m × 1200 m square area centered on each use location. (C) We generated use locations 2 and 3 within a 1200 m × 1200 m square area centered on the previous use location and generated random locations within a 600 m × 600 m square area centered on each use location. (D) We generated use locations 2 and 3 within a 1200 m × 1200 m square area centered on the previous use location and generated random locations within a 1200 m × 1200 m square area centered on each use location. When analyzing the data using discrete choice, the color-coordinated use and random locations were stratified and defined the choice sets. For logistic regression and Maxent, all random locations were evaluated with each use location and defined the choice sets.

calculated the mean and standard error of slope and intercept coefficients from linear regressions of the 100 models. Methods that produced unbiased estimates of the known probability distribution would have mean slope and intercept values from linear-regression models of 1.0 and 0.0, respectively. Estimates of the known distribution were considered acceptable when 95% confidence intervals

around mean slope and intercept coefficients contained 1.0 or 0.0, respectively.

For discrete choice and logistic regression, we scaled and centered distance measures for all resources across the landscape using: $DIST_{scaled} = (OBS - \mu)/SD$. In this equation, $DIST_{scaled}$ were scaled distance measures used in the models, OBS were distances

of resource units to respective features, and μ and SD were the mean and standard deviation of respective distance measures in the set of data, respectively. After scaling, each continuous distance measure had a mean of 0.0 and variance of 1.0. Maxent used ($DIST_{clamped} = (OBS - MIN_{obs})/(MAX_{obs} - MIN_{obs})$) to "clamp" or reduce the range of continuous variables to have a minimum value of 0.0 and a maximum value of 1.0 for all locations within the set of data. In this equation, $DIST_{clamped}$ are clamped distance measures used in the models, OBS was the observed distance of location to respective feature, and MIN_{obs} and MAX_{obs} were minimum and maximum values for distance to respective features in the simulated set of data. We used Eq. (5) (Appendix A.2) to calculate Maxent probabilities of selection across the landscape and assigned a value of 0.0 or 1.0 to all resources with distances < or > what was observed in the simulated set of data, respectively. For example, if a continuous variable had a range of −100 to 100 and a particular set of simulated data had a range of −90 to 80, we assigned a value of 0.0 or 1.0 to all pixels across the landscape with a distance measure <−90 or >80, respectively, when computing probabilities of resource selection across the landscape. We used linear-regression procedures outlined for discrete choice and logistic-regression to determine the ability of Maxent to estimate the known underlying probability distribution.

We also measured model performance by comparing AUC scores produced in the 100 simulations. We used 75% of use locations along with the sample of random locations drawn randomly without replacement from the area defined to be available to produce models for receiver operating characteristic (ROC) analyses. We used the remaining 25% of use locations as a test sample to calculate AUC scores for discrete-choice and logistic-regression models (Appendix A.1(5)). To produce AUC scores for Maxent, we set the "test sample" size to 25% and AUC values were included in the output as a standard function of the program. We produced boxplots to display the average and range in distribution of AUC values (Appendix A.1(6)). We observed trends in AUC values to ascertain improvements in fit of models to test data across all levels of use and availability and across changes in number of animals and random locations used to produce each model.

We also compared the ability of each method to reproduce coefficients of the known model. Differences in how continuous variables were standardized (scaled and centered, clamped) and representation of the categorical variable in each model (intercept, reference category, estimation of all levels), however, precluded an in-depth comparison of all βs. Estimates of coefficients of the landcover variable produced in analyses led to similar conclusions as other tests of performance so we did not report these results.

3. Results

3.1. Ability of statistical methods to estimate the known probability distribution

Increasing the number of random locations used in analyses from 5 random locations per use location to 10,000 random locations (total) typically resulted in increased precision and accuracy of estimates of the known probability distribution for all statistical methods, but within each method, these estimates were statistically indistinguishable. We provided plots of average slope coefficients and approximate 95% confidence intervals for all three statistical methods, across various levels of use and availability, and for sample sizes of 10 and 80 animals with 25 use locations per animal and 5 random locations per use location (Fig. 2); results from analyses with 20 and 40 animals fell between those two extremes. We did not include a plot of intercept coefficients as the range of all values was between −0.0001 and 0.00005.

The precision of discrete-choice estimates of the known probability distribution increased (smaller standard errors) as we increased sample size from 10 to 80 animals (Fig. 2). The accuracy of estimates of the known probability distribution (bias), however, was not affected by the number of animals included in discretechoice analyses. Estimates of the known probability distribution produced by discrete-choice models were usually unbiased and precise when area defined to be available was no larger than the area in which use occurred. Increasing the number of random locations from 5 random locations per use location to 10,000 random locations total had little or no effect on the precision or accuracy of estimates of the known probability distribution produced by discrete-choice models. Across all sample sizes, estimates of the known probability distribution produced by discrete-choice models were usually more precise and accurate than logistic-regression or Maxent when the area defined to be available was at a small scale.

The precision of estimates of the known probability distribution produced by logistic-regression models increased as the number of animals included in the analysis increased (Fig. 2). The accuracy of these estimates, however, declined as we increased the number of animals included in the analysis or decreased the area defined to be available. When we generated use and random locations at larger scales ($\geq$1200 m), estimates of the known probability distribution produced by logistic-regression models were unbiased and precise (Fig. 2). Estimates of the known probability distribution produced in logistic-regression analyses were unbiased and most precise when levels of use and availability were equal and the ratio between the number of random locations and use locations used in the analysis increased from 5:1 to 40:1.

The number of animals and random locations used in analyses and definition of availability affected the accuracy of estimates of the known probability distribution produced by Maxent (Fig. 2). Similar to other methods, the precision of estimates of the known probability distribution produced by Maxent increased as we increased the number of animals and random locations and as the level of availability increased. The accuracy of these estimates, however, decreased as the number of animals increased and the level of availability used in the analysis decreased. Maxent produced unbiased and precise estimates of the known probability distribution when the ratio between level of use and availability was smallest (300 m: DNWR, Fig. 2).

3.2. Effect of level of use, level of availability, and sample size on ROC analyses

We examined boxplots of AUC scores for all three statistical methods, across all levels of use and availability, and for all sample sizes. We observed a trend in the variability and range of AUC scores across numbers of animals used in analyses and levels of availability and provided plots of AUC scores from analyses with 10 or 80 animals, 5 random locations per use location, and across various levels of use and availability (Fig. 3). Increasing the number of animals and random locations did not affect mean values of AUC scores, but did influence the variability in AUC scores. Variability in AUC scores was greater at smaller sample sizes (10 animals) than at larger sample sizes (80 animals). As the ratio between level of use and availability decreased, AUC scores increased for all statistical methods tested (Fig. 3). We found no difference in AUC scores across statistical methods tested.

4. Discussion

Our results highlight the critical importance of scale in assigning use and availability in studies of resource selection (Boyce, 2006). The area in which we generated use and random locations
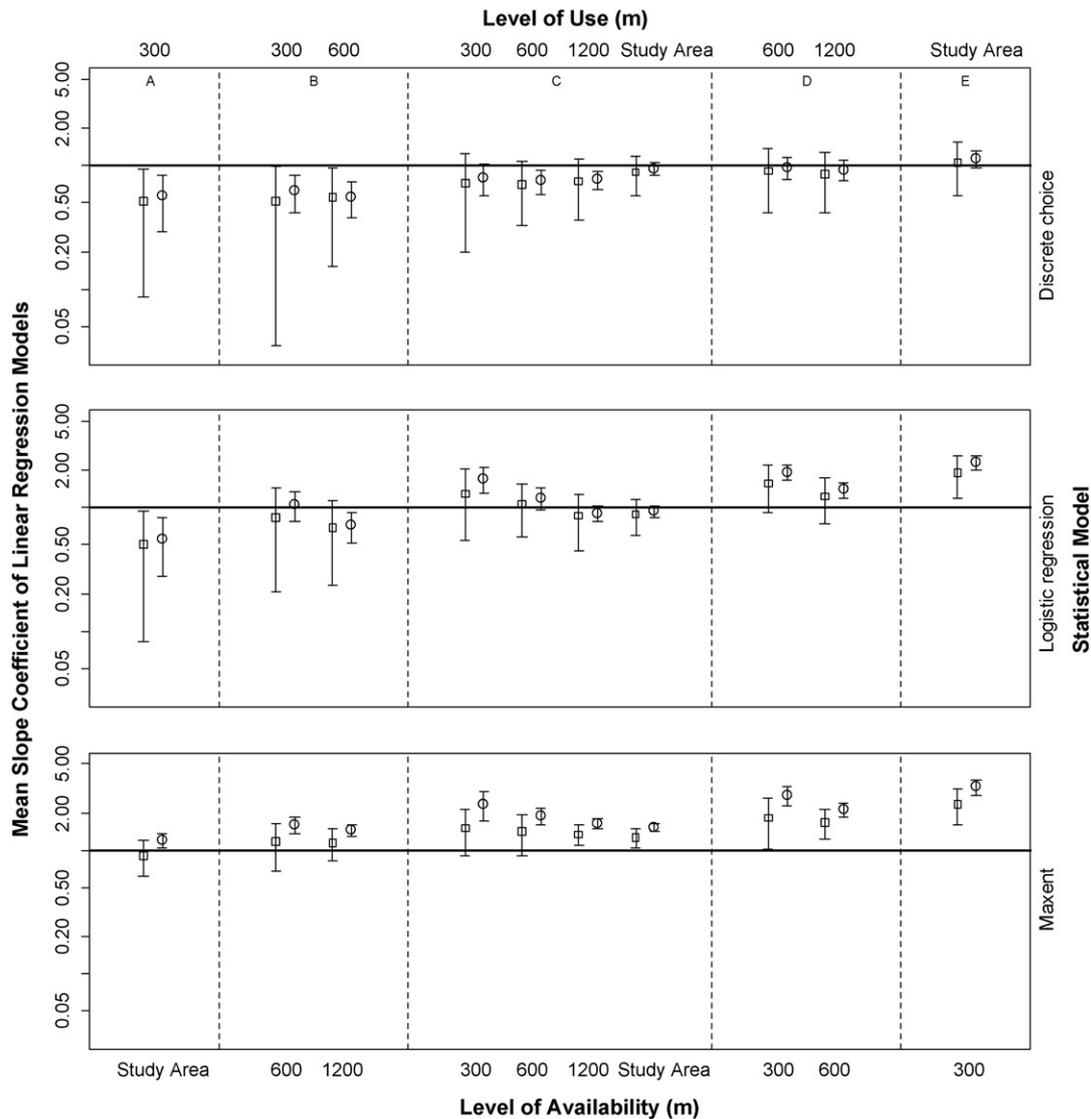
Fig. 2. Mean slope coefficient of linear regression models with known probabilities predicted by estimated probabilities of resource selection and approximate 95% confidence intervals ($\bar{X} \pm 1.96SE$) for discrete-choice, logistic-regression, and Maxent models at sample sizes of 10 (□) or 80 (○) animals with 25 use locations per animal and 5 random locations per use location. Plots are grouped where the ratio between area in which we generated use locations (level of use) and area defined as available was smallest (300 m: study area (A)), 1:2 (B), 1:1 (C), 2:1 (D), and largest (study area: 300 m (E)).

affected all statistical methods, but the effect differed by method. For discrete-choice models, if selection of resources occurs within a localized area (i.e., within a home range) the area used to define the choice set should be no larger than the area in which use actually occurred. To determine the area in which use occurred, many studies have used displacement distances between radio-telemetry observations (Arthur et al., 1996; Cooper and Millspaugh, 1999). When using discrete choice to model resource selection, our findings support definitions of availability used in these studies, such as the definition used by Cooper and Millspaugh (1999). Cooper and Millspaugh (1999) used a circular area with radius equal to one-half of the average displacement distance centered one-quarter day's walk from a previous telemetry location in the direction of a known bed site to define what was available to elk (*Cervus elaphus*) in South Dakota. Arthur et al. (1996) used a circular area centered on the previous location with radius equal to the distance polar bears (*Ursus maritimus*) in the Chukchi and Bearing Seas were "likely" to travel within 3 or 6 days dependent on the number of days between collections of use locations. Both of these, and other studies (McCracken et al., 1998), suggest centering the choice set

for subsequent use locations on or in the direction of the previous use location. Such definitions are logical if one assumes the animal has no *a priori* knowledge of the area in which they select use locations, but in reality, animals are typically well acquainted with their surroundings. A potential problem of such definitions for the choice set is that the area in which the choice set is defined may not contain the subsequent use location. For example, Arthur et al. (1996) reported that on 2630 occasions (<2%), bears traveled a linear distance ≥ what was defined to be available, which would result in these use locations falling outside the area in which the choice sets were defined. Similar problems could also arise if one used the definition of choice set suggested by Cooper and Millspaugh (1999) and McCracken et al. (1998). Another potential problem of such definitions is the uncertainty in the direction the animal approached the subsequent use location if one does not collect observations on a regular basis (∼28 h) as did Cooper and Millspaugh (1999). For these reasons, we centered the choice set for the current location on the actual location rather than on or in the direction of the previous one. So long as the area defined to be available does not contain resources that are inaccessible due to excessive distance from pre-
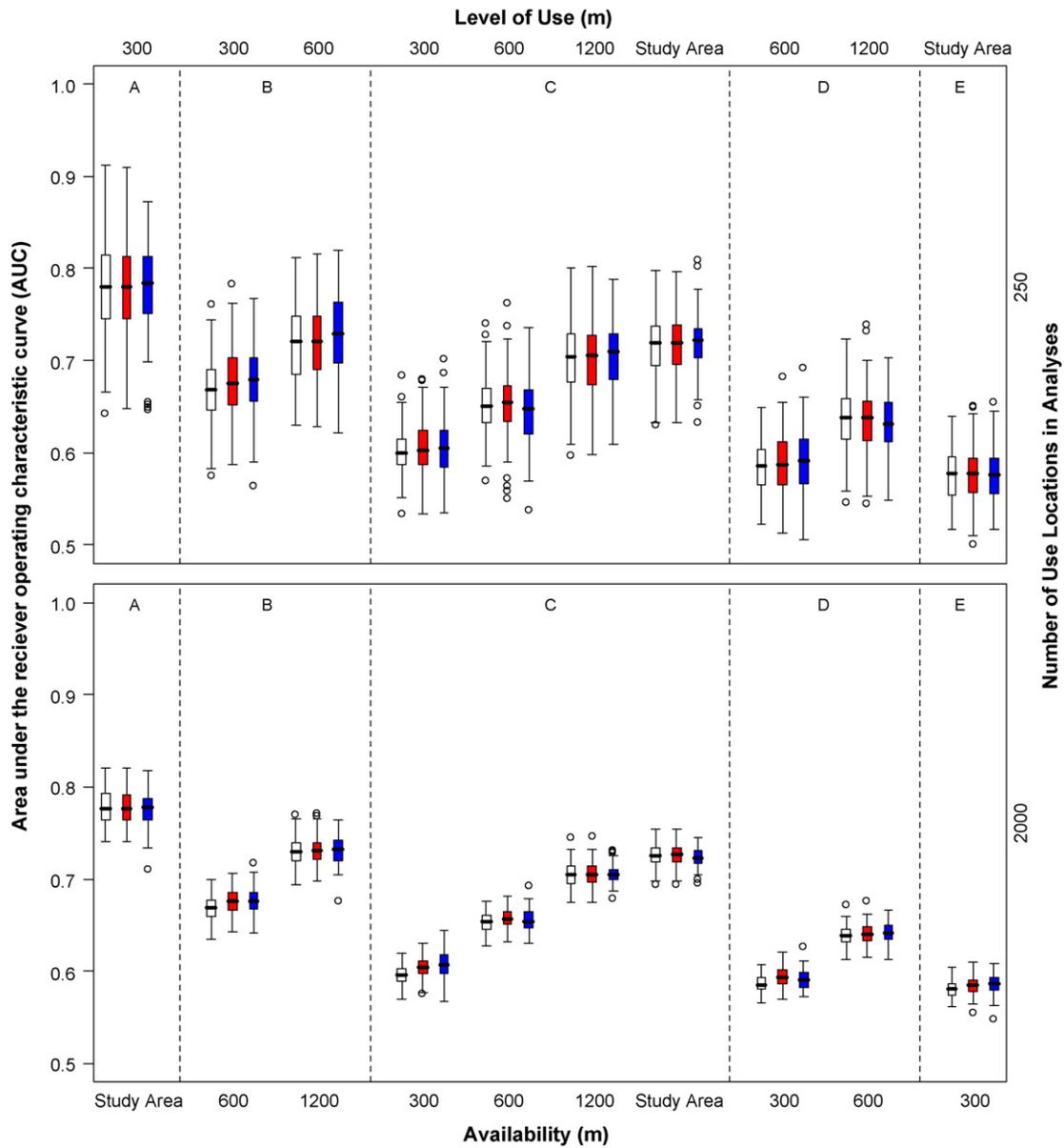
Fig. 3. Area under receiver operating characteristic curve (AUC) scores for discrete-choice (white), logistic-regression (red), and MAXENT models (blue) at sample sizes of 10 or 80 animals with 25 use locations per animal and 5 random locations per use location. Plots are grouped where the ratio between area in which we generated use locations (level of use) and area defined as available was smallest (300 m: study area (A)), 1:2 (B), 1:1 (C), 2:1 (D), and largest (study area: 300 m (E)). Boxes encompass central 50% of observations and length of whiskers = 1.5 × interquartile of the box.

vious location or other prohibitive factor, this definition should be appropriate and would ensure entire choice sets were contained in the sample.

When using discrete-choice models to analyze data collected in areas, such as our study area, where average size of habitat patches was relatively small ($\bar{X} = 10.6$ ha) with few or no avoided areas, selecting random locations within a small area centered on each use location was not problematic. In areas where patch sizes are much larger than those in our study area, however, one could lose important effects of variables if selection of random locations occurs within small, localized areas around use locations that do not include areas avoided by the animal. Multi-level analyses of resource selection can help avoid such pitfalls in these situations (Johnson, 1980; Aebischer et al., 1993; McClean et al., 1998; Thompson and McGarigal, 2002; D'Eon and Serrouya, 2005). Our results suggest, when using discrete choice, one could use displacement distances between successive use locations to determine

factors that influence selection at a larger, macro-habitat scale and potentially a smaller scale (one-eighth- or one-quarter-days walk) to determine factors that influence selection at a smaller, micro-habitat scale.

Logistic-regression and Maxent modeling resulted in accurate and precise estimates of the known probability distribution when the study area was defined to be available. Such definitions of availability are politically rather than biologically determined boundaries and can be problematic if the study area is large or the research animal is not very mobile (e.g., turtle). If the area defined to be available includes resources outside the range selected by the animal, resources may not be accessible due to distance, competition, or other preclusive scenarios (Johnson, 1980; Aebischer et al., 1993; Keating and Cherry, 2004). A notable difference between these two techniques was that estimates of the known probability distribution were most accurate and precise for logistic-regression when the levels of use and availability both occurred on a larger

scale (≥1200 m) and for Maxent when use occurred within a local-ized area (e.g., within a home range) and availability was defined at a larger scale (study area). The fact that Maxent performed best under these conditions is reasonable given its successful applica-tion in the field of species distribution mapping (Phillips et al., 2006). In many species distribution studies, occurrence or use loca-tions are obtained from specimens in natural history museums and herbaria and "background" or random locations are typically drawn from across the study area to define the space in which use occurred (Ponder et al., 2001; Hoffman, 2008). Occurrence data in many of these studies tend to be clustered groups of locations in various regions of the study area (e.g., Fig. 2 in Phillips et al., 2006), which could be related to differences in intensity of sam-pling efforts across the study areas (Ponder et al., 2001). In these situations, the distribution of use locations (occurrence) and ran-dom locations (background) would be similar to our study when use occurred within a localized area (300 m) and availability was defined at the study-area level.

Previous studies of statistical techniques used in analyses of resource selection suggest the number of animals included in the sample should be ≥20 with 50 locations/animal (Alldrege and Ratti, 1986, 1992; Leban et al., 2001). Resource selection models evalu-ated in these studies, however, were nonparametric or multivariate approaches with 1 categorical variable (landcover) included in the models. Our multivariate approaches that included categorical and continuous variables showed that increasing the number of ani-mals used in analyses had little effect or resulted in slightly less accurate estimates of the known probability distribution. The pre-cision of estimates, however, improved as the number animals increased, which was similar to results reported by others (Manly, 2002; Manly et al., 2002). It was surprising that the ability to accu-rately estimate resource selection functions with samples as small as 10 animals with 25 locations/animal was observed in this study; this could be advantageous when analyzing data obtained from a small population of animals or when budget or time constraints prohibit more intensive sampling efforts. Performance in individual study areas with different underlying features, however, is neces-sary before our results are applied elsewhere.

Similar to McFadden's (1978) findings, increasing the number of random locations used in discrete-choice analyses had little effect on estimates of the known probability distribution when the area in which use locations were generated was greater than the area in which random locations were generated. When levels of use and availability were misclassified (level of availability > level of use), results of discrete-choice analyses improved slightly by increasing numbers of use and random locations, but the improvement did not overcome the misclassification error. Increasing the number of random locations used in Maxent and logistic-regression analy-ses, however, tended to improve the accuracy of estimates of the known probability distribution. A potential problem of using an increased number of random locations in analyses is contamina-tion (resources actually used, but classified as random locations; Keating and Cherry, 2004). Contamination may or may not affect the results of analyses depending upon how robust the statistical method is and the level of contamination in the sample (Keating and Cherry, 2004; Johnson et al., 2006).

Interpretation of a statistical method's performance varied depending on the method used for evaluation. The AUC analy-ses often resulted in different conclusions of model performance than comparisons of estimated and known probability distribu-tions. A generally accepted guideline for determining acceptability of models is that models with good predictive ability have AUC scores >0.75 (Elith, 2002; Phillips and Dudík, 2008). According to this criterion, all models had good predictive ability when selec-tion of use locations occurred within a smaller area and availability was defined at the study-area level (Fig. 3). Results of AUC analy-ses consistently ranked models produced by Maxent with high and low predictive abilities accurately. As random sampling extended to the "study area" and the area of selection decreased, Max-ent performed better than other methods. In these situations, it was more likely that unused resources were included in the sam-ple of random locations (decreased contamination), the disparity between probabilities of selection between used and random loca-tions increased, and thus higher AUC scores. We found, however, AUC scores increased as the area defined to be available increased, which often resulted in an inaccurate measure of performance for discrete-choice and logistic-regression models. Discrete-choice models performed best when the area of use was ≥ area defined to be available, which likely resulted in similar probabilities of selection between used and random locations, and thus lower AUC scores. As AUC score are a measure of the models ability to rank a use location higher than a random one, it was not a reliable mea-sure of performance for the more traditional RSF models (Boyce et al., 2002; Termansen et al., 2006; Austin, 2007; Lobo et al., 2008); calibration curves comparing estimated and known probability distributions were the most telling evaluation of each statistical method. A more complete assessment of AUC methods is needed to determine their utility in evaluating RSFs.

5. Conclusion

Discrete-choice models consistently produced the most accu-rate and precise estimates of the known probability distribution in our study area when levels of use and availability were specified correctly, regardless of the number of animals or random locations included in analyses. When using discrete-choice models to esti-mate resource selection, the area defined as available should be restricted to a small area in which resources were selected. If the level in which use occurred is unknown, one should err on the side of defining a smaller area of availability. Distribution and size of patches of landcover, however, could also affect the performance of statistical methods used in resource selection studies so further studies using simulated data over different landscapes and with dif-ferent data types are needed to confirm our results. We urge caution when using ROC analyses as a measure of model performance in resource selection studies, especially when standard statistical pro-cedures such as logistic-regression or discrete choice are used. Our study offers a template and guidance for others interested in assess-ing techniques for modeling resource selection for their study site; the performance of these and other statistical methods may vary, however, when applied to study areas with different underlying features.

Appendix A.

A.1. Description of functions, packages, and additional settings in Program R (R Development Core Team, 2006) used to select use and random locations, produce models, evaluate model performance, and create figures

Reference	Function	Package	Additional Settings
1	rmultinom	stats	NA
2	sample	base	NA
3	clogit[a]	survival[b]	NA
4	glm	stats	logit[c]
5	performance	ROCR	auc[d]
6	boxplot	graphics	NA

[a] The "clogit" function is a wrapper for the Cox Proportional Hazards regres-sion that estimates a logistic-regression model by maximizing the exact conditional likelihood (R Development Core Team, 2006).

[b] Lumley (2006).

[c] Link used to produce logistic-regression models.

[d] Performance measure setting used to produce AUC scores for discrete-choice and logistic-regression models.

A.2. Overview of statistical models

The standard equation for the discrete-choice model (McDonald et al., 2006) was:

$$\hat{P}(i) = \frac{\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})}{\sum_{k \in \{U' \cup A\}} \exp(\beta_1 x_{k1} + \beta_2 x_{k2} + \cdots + \beta_p x_{kp})}, \tag{1}$$

where $\hat{P}(i)$ was an estimate of the probability of selection for resource unit i; x_{i1}–x_{ip} were characteristics of resource units; β_1–β_p were coefficients of respective characteristics; U' were the set of indices for unique used units; and A were the set of indices for units in the random sample of units from the choice set.

The standard equation for the logistic-regression model was:

$$\hat{P}(i) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})}, \tag{2}$$

where $\hat{P}(i)$ was an estimate of probability of use for resource unit i; x_{i1}–x_{ip} were characteristics of resource units (landcover and distance measures); and β_1–β_p were coefficients of respective characteristics.

The standard equation for a Maxent model (Phillips et al., 2006; Phillips and Dudík, 2008) was:

$$q_\lambda(x) = \frac{\exp\left(\sum_{i=1}^{p} \lambda_i f_i(x)\right)}{Z_\lambda}, \tag{3}$$

where $q_\lambda(x)$ was an estimate of $(\hat{P}[y = 1|x])$ or probability of selection for resource unit x; λ were coefficients of the model; $f(x)$ were characteristics of the resource units; and $Z_\lambda = \sum_{x \in x} x^{e^{\lambda * f|(x)}}$ was a normalizing constant used to ensure probabilities sum to 1.0. Where we only modeled linear terms, $\lambda_i f(x_i)$ was simply $\lambda_1 x_{i1} + \lambda_2 x_{i2} + \cdots + \lambda_p x_{ip}$. To obtain Maxent estimates of the conditional probability of selection for all resource units within the study area (Phillips and Dudík, 2008), we used:

$$Q(y = 1|x) = \frac{e^H q_\lambda(x)}{1 + e^H q_\lambda(x)}, \tag{4}$$

where q_λ was the maximum entropy estimate of the true distribution of selection; and H was the entropy of the Maxent model, q_λ (Phillips and Dudík, 2008).

We present Eq. (4) with a similar notation used in Eqs. (1) and (2) which resulted in:

$$\hat{P}(i) = \frac{\exp(H)(\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})/Z_\beta)}{1 + \exp(H)(\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})/Z_\beta)} \tag{5}$$

where $\hat{P}(i)$ was the conditional probability of resource unit i being selected; x_{ip} were the characteristics of resource unit i; and β_1–β_p were coefficients of respective characteristics.

Eq. (5) is the logistic format of the Maxent model (Phillips and Dudík, 2008), which is similar to the logistic-regression model (Eq. (2)); however, the Maxent model contains a parameter estimate for all levels of the categorical variable (landcover) and an intercept estimated by the entropy of the model.

References

Aebischer, N.J., Robertson, P.A., Kenward, R.E., 1993. Compositional analysis of habitat use from animal radio-tracking data. Ecology 74, 1313–1325.

Alldrege, J.R., Ratti, J.T., 1986. Comparison of some statistical techniques for analysis of resource selection. J. Wildl. Manage. 50, 157–165.

Alldrege, J.R., Ratti, J.T., 1992. Further comparison of some statistical techniques for analysis of resource selection. J. Wildl. Manage. 56, 1–9.

Arthur, S.M., Manly, B.E., McDonald, L.L., Garner, G.W., 1996. Assessing habitat selection when availability changes. Ecology 77, 215–227.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Model. 157, 101–118.

Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol. Model. 200, 1–19.

Baasch, D.M., 2008. Resource selection by white-tailed deer, mule deer, and elk in Nebraska. Dissertation. University of Nebraska, Lincoln, USA.

Berger, A.L., Della Pietra, S.A., Della Pietra, V.J., 1996. A maximum entropy approach to natural language processing. Comput. Linguist. 22, 39–71.

Berger, U., Wagner, G., Wolff, W.F., 1999. Virtual biologists observe virtual grasshoppers: an assessment of different mobility parameters for the analysis of movement patterns. Ecol. Model. 115, 119–127.

Boyce, M.S., 2006. Scale for resource selection functions. Divers. Distrib. 12, 269–276.

Boyce, M.S., Mao, J.S., Merrill, E.H., Fortin, D., Turner, M.G., Fryxell, J., Turchin, P., 2003. Scale and heterogeneity in resource selection by elk in Yellowstone National Park. Ecoscience 10, 421–431.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. Ecol. Model. 157, 281–300.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. Lett. 30, 1145–1159.

Buskirk, S.W., Millspaugh, J.J., 2006. Metrics of use and availability in studies of resource selection. J. Wildl. Manage. 70, 358–366.

Collins, M., Schapire, R.E., Singer, Y., 2002. Logistic regression, AdaBoost and Bregman distances. Mach. Learn. 48, 253–285.

Cooper, A.B., Millspaugh, J.J., 1999. The application of discrete choice models to wildlife resource selection studies. Ecology 80, 566–575.

D'Eon, R.G., Serrouya, R., 2005. Mule deer seasonal movements and multiscale resource selection using global positioning system radiotelemetry. J. Mammal. 86, 736–744.

Della Pietra, S., Della Pietra, V., Lafferty, J., 1997. Inducing features of random fields. IEEE Trans. Pattern. Anal. Mach. 19, 1–13.

Dudík, M., Phillips, S.J., Schapire, R.E., 2004. Performance guarantees for regularized maximum entropy density estimation. In: Proceedings of the 17th Annual Conference on Computational Learning Theory. ACM Press, New York, NY, pp. 655–662.

Elith, J., 2002. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson, S., Burgman, M. (Eds.), Quantitative Methods for Conservation Biology. Springer, New York, NY, USA, pp. 39–58.

Erickson, W.P., McDonald, T.L., Gerow, K.G., Howlin, S., Kern, J.W., 2001. Statistical issues in resource selection studies with radio-marked animals. In: Millspaugh, J., Marzluff, J. (Eds.), Radio Tracking and Animal Populations. Academic Press, San Diego, CA, USA, pp. 211–242.

ESRI, 2006. ArcGIS 9. 2 Geographic Information System. Environmental Systems Research Institute Inc., Redlands, CA, USA.

Fawcett, T., 2005. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874.

Garshelis, D.L., 2000. Delusions in habitat evaluation: measuring use, selection, and importance. In: Boitani, L., Fuller, T. (Eds.), Research Techniques in Animal Ecology: Controversies and Consequences. Columbia University, New York, NY, USA, pp. 111–164.

Hanley, J., McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29–36.

Hirzel, A.H., Helfer, V., Metral, F., 2001. Assessing habitat suitability models with a virtual species. Ecol. Model. 145, 111–121.

Hoffman, J.D., 2008. Evaluation and application of predictive habitat modeling in ecology. Dissertation. University of Nebraska, Lincoln, USA.

Johnson, C.J., Nielsen, S.E., Merrill, E.H., McDonald, T.L., Boyce, M.S., 2006. Resource selection functions based on use-availability data: theoretical motivation and evaluation methods. J. Wildl. Manage. 70, 347–357.

Johnson, D.H., 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 61, 65–71.

Keating, K.A., Cherry, S., 2004. Use and interpretation of logistic regression in habitat-selection studies. J. Wildl. Manage. 68, 774–789.

Leban, F.A., Wisdom, M.J., Garton, E.O., Johnson, B.K., Kie, J.G., 2001. Effect of sample size on the performance of resource selection analysis. In: Millspaugh, J.J., Marzluff, J.M. (Eds.), Radio Tracking and Animal Populations. Academic Press, San Diego, CA, USA, pp. 291–307.

Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecol. Biogeogr. 17, 145–151.

Lumley, T., 2006. S original by Terry Therneau and ported by Thomas Lumley. Survival: Survival Analysis, Including Penalised Likelihood. R package version 2.31.

MacKenzie, D.J., 2006. Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. J. Wildl. Manage. 70, 367–374.

Manly, B.F.J., 2002. Estimating a resource selection function with line transect sampling. J. Appl. Math. Decis. Sci. 6, 213–228.

Manly, B.F.J., McDonald, L.L., Thomas, D.L., 1993. Resource Selection by Animals: Statistical Design and Analysis for Field Studies. Chapman and Hall, London, United Kingdom, p. 192.

Manly, B.F.J., McDonald, L.L., Thomas, D.L., McDonald, T.L., Erickson, W.P., 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies, second edition. Kluwer Academic Publishers, Boston, MA, USA, p. 240.

McClean, S.A., Rumble, M.A., King, R.M., Baker, W.L., 1998. Evaluation of resource selection methods with different definitions of availability. J. Wildl. Manage. 62, 793–801.

McCracken, M.L., Manly, B.F.J., Heyden, M.V., 1998. The use of discrete-choice models for evaluating resource selection. J. Agric. Biol. Environ. Stat. 3, 268–279.

McDonald, T.L., Manly, B.F.J., Nielson, R.M., Diller, L.V., 2006. Discrete-choice modeling in wildlife studies exemplified by northern spotted owl nighttime habitat selection. J. Wildl. Manage. 70, 375–383.

McFadden, D., 1978. Modeling the choice of residential location. In: Karlquist, A. (Ed.), Spatial Interaction Theory and Planning Models. North Holland Publishing Company, Amsterdam, The Netherlands, pp. 79–96.

Neu, C.W., Beyers, C.R., Peek, J.M., 1974. A technique for analysis of utilization—availability data. J. Wildl. Manage. 38, 541–545.

Owen, M., 1972. Some factors affecting food intake and selection in white-fronted geese. J. Anim. Ecol. 41, 79–92.

Özesmi, U., Tan, C.O., Özesmi, S.L., Robertson, R.J., 2006. Generalizability of artificial neural network models in ecological applications: predicting nest occurrence and breeding success of the red-winged blackbird *Agelaius phoeniceus*. Ecol. Model. 195, 94–104.

Park, S.H., Goo, J.M., Jo, C.-H., 2004. Receiver operating characteristic (ROC) curve: practical review for radiologists. Kor. J. Radiol. 5, 11–18.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.

Phillips, S.J., Dudík, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling. In: Proceedings of the 21st International Conference on Machine Learning. ACM Press, New York, NY, pp. 655–662.

Phillips, S.J., Dudík, M., Schapire, R.E., 2005. MAXENT Software for Species Distribution Modeling., http://www.cs.princeton.edu/~schapire/maxent/.

Ponder, W.F., Carter, G.A., Flemons, P., Chapman, R.R., 2001. Evaluation of museum collection data for use in biodiversity assessment. Conserv. Biol. 15, 648–657.

Quade, D., 1979. Using weighted rankings in the analysis of complete blocks with additive block effects. J. Am. Stat. Assoc. 74, 680–683.

R Development Core Team, 2006. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, http://www.R-project.org.

Termansen, M., McClean, C.J., Preston, C.D., 2006. The use of genetic algorithms and Bayesian classification to model species distributions. Ecol. Model. 192, 410–424.

Thomas, D.L., Johnson, D., Griffith, B., 2006. A Bayesian random effects discrete-choice model for resource selection: population level selection inference. J. Wildl. Manage. 70, 404–412.

Thompson, C.M., McGarigal, K., 2002. The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). Land Ecol. 17, 569–586.

Tyre, A.J., Possingham, H.P., Lindenmayer, D.B., 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters. Ecol. Appl. 11, 1722–1737.