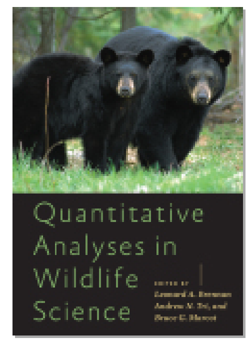
Quantitative Analyses in Wildlife Science

Brennan, Leonard A., Tri, Andrew N., Marcot, Bruce G.



Published by Johns Hopkins University Press

PART IV ANALYSIS OF SPATIALLY BASED DATA ON ANIMALS AND RESOURCES

This page intentionally left blank

12 Resource Selection Analysis

Julianna M. A. Jenkins,
Damon B. Lesmeister,
and Raymond J. Davis

Different taxa use resources differently across a single day, season, or life cycle and these dependencies must be quantified in order to understand management.
—Mathewson and Morrison (2015:4)

Introduction

The process of an animal selecting resources involves a series of behavioral choices. Understanding these behaviors is a foundational and cross-cutting theme in wildlife research and management. Resource selection analyses (RSAs) represent a broad class of statistical models for identifying underlying environmental correlates of animal resource selection and space use patterns. Many other chapters of this book describe analyses that can be encompassed by the general RSA definition. RSAs encompass several categories of habitat analyses: methods focused on testing for disproportionate use of habitat components, often referred to as "use-availability" or "presence-absence" models (Manly et al. 2002; Johnson et al. 2006), and "presence-only" models (McDonald 2013; Warton and Aarts 2013). RSAs have been used to improve our basic understanding of wildlife ecology (e.g., hypothesis testing, ecological niche models or species distribution models), to predict animal space use, to inform habitat management and conservation (e.g., critical habitat delineation), and to assess impacts of environmental change on wildlife (e.g., wildlife spatial response to wildfire).

Volumes have been written on the subject of RSAs and despite—or because of—this, choosing the most appropriate method of analysis for any given study question can be an overwhelming task. The goal of this chapter is to dispel the confusion and to present a basic introduction to resource selection concepts and an overview of common RSA methods used by wildlife professionals. This is not meant to be a comprehensive review, but rather an introductory guide. Other reviews and primary sources are referenced throughout the chapter, which we strongly suggest reading for a broader understanding of RSA theory and practice. Manly et al. (2002), the March 2006 special section of the *Journal of Wildlife Management* (Strickland and McDonald 2006), and the November 2013 special feature in the *Journal of Animal Ecology* (Mcdonald et al. 2013) provide more extensive discussions and reviews of specific RSAs.

Foundational Concepts

Resources and Resource Units

Habitat is a comprehensive term describing an area that encompasses the necessary combination of environmental conditions and resources that promote occupancy, reproduction, and survival of a species

(Morrison et al. 2006). Therefore, wildlife biologists interested in understanding habitat are typically focused on the identification, availability, and relative importance of resources. The definition of a *resource* in wildlife biology is broad and includes: matter taken up by an animal (e.g., food items), objects with which animals associate (e.g., nest tree), and conditions that influence the use of places and ultimately affect fitness (e.g., vegetation cover type; Buskirk and Millspaugh 2006). *Resource units* are quantifiable items or areas that can be observed as used (or not) by an animal, sometimes also called sample units (Lele et al. 2013; Rota et al. 2013). We focus our discussion on resource units comprised of spatial areas (e.g., quadrats) since these are most common in wildlife ecology. *Available resource units* are those units that are accessible to animals during a period of interest (Johnson 1980). *Used resources* are by definition a subset of available resources that are encountered and utilized, while *unused resource units* can be defined as either available or unavailable based upon study designs and assumptions.

Resource units can be described by a single attribute (e.g., canopy cover) or multiple attributes (e.g., land unit's slope and soil type) that differ among units. The method used to describe resource units affects the appropriate analysis, model interpretation, and study costs. Improvements in remotely sensed environmental attributes and geographic information systems (GIS) technology have greatly expanded our ability to accurately describe some resource units over large spatial and temporal scales, but many studies still rely upon detailed ground surveys. Although acquiring and classifying resource data into broad categories may be quicker and more cost effective than producing detailed measurements in the field, general classifications may miss important ecological mechanisms of selection. For example, height of groundcover may be the most important variable for a generalist bird's selection of a nest site, and broad vegetation classifications such as grassland versus forest would not capture the importance of groundcover. However, if the purpose of an analysis is to determine how animals respond to a general management action or disturbance, such as a forest harvest, then categorizing resource units as either harvested or unharvested, or by their distance to a harvest, may be sufficient.

The distribution of available resources describes the variation of resource types in environmental space (Lele et al. 2013). For example, when resource units are described by a single categorical attribute, say cover type, the available resource distribution represents the proportion of available units in each cover type category within some given area. When a resource is used disproportionately to its availability we infer selection (Johnson 1980). A statistical model used to estimate the probability of selecting a resource unit as a function of resource attributes is often called a *resource selection function* (RSF) or *resource selection probability function* (RSPF; Manly et al. 2002; Lele et al. 2013).

Availability

Defining what resources are available to individuals or a population is a primary concern for anyone interested in resource selection (Lele et al. 2013; Manly et al. 2002). By definition, available resources are assumed to be accessible to focal animals during a period of interest (Johnson 1980). Available resources can be completely accounted for, or a random sample of available units can be sampled from the available extent (Manly et al. 2002). In practice accessibility is rarely empirically studied partly because the focus of most studies is on the used resources, and quantifying availability is often constrained by logistics. Used units are typically compared with units occurring (existence or abundance) in some portion of an animal's environment. Boundaries of available resource occurrence are based upon assumptions from animal movement studies (e.g., home range polygons), management area boundaries, and mapping extents (Elith et al. 2011), or they are limited by budget or personnel. Consequently, when available resource units and distributions are summarized, results likely represent a greater diversity of attributes and quantity of resource units than what was truly accessible to the focal animal(s), and the selectivity metrics calculated can be invalid or applicable only to the study area investigated (Buskirk and Millspaugh 2006). The determination and underlying assumptions of available resources are important factors in determining how well a model actually represents the population of interest, regardless of the statistical method.

Use

When food items are resource units, the definition of use is straightforward: a used unit is one that is eaten. When spatial areas are considered resource units, interpretations of "use" are more variable. The presence of an animal on a resource unit typically corresponds to "use"; however, spatial units can be selected more than once and areas may be visited at different rates depending upon the size of resource units, timing of sampling, and animal's behavior during an observation. Use can be defined as binary, such as used versus unused or used versus available, or use can be defined by a measure of use intensity. In many studies, unused areas are difficult to determine because observations are snapshots in time. The method of observation affects how use is measured. When individual animals are not identified, use could be inferred as the presence or abundance of a species or individuals in a sampled area or unit and measured at each location or in a buffered area or plot containing each location. For example, Neu et al. (1974) recorded use as the number of moose tracks in four plots with variable burn history over seven time periods. When individual animals are identified with repeated observations of use generated for each animal, the location of individual relocations, clusters of relocations, or the area encompassing all relocations can be used to define use. Aebischer et al. (1993) defined use as the proportion of resources within an individual animal's entire home range area, while Millspaugh et al. (2006) divided each home range into a grid and defined use as the mean utilization distribution (UD) value within each grid cell. Rate of use can also vary by the time of year and by animal behavior. For example, resources selected for foraging in spring may be vastly different from resources selected for nesting or from resources selected in winter. Some analyses can be designed to accommodate these dynamics. The type of use data collected is a primary determinant of which RSA is appropriate.

Scales of Selection and Use

Resource selection was described by Johnson (1980) as a hierarchical process, whereby each order of selection is conditional on a selection made at preceding levels. *First-order selection* is defined at the level of the physical or geographic range of a species; *second-order selection* is defined as selection of a use-area conditional on the species range (e.g., home range); *third-order selection* is defined as use of a component within the second order resource (e.g., cover type); and *fourth-order selection* refers to use of a particular resource within the third order resource (e.g., nest tree; Johnson 1980). The resources important to selection may change between or among orders of selection. For example, the geographic range of a species (first-order) may be limited by climatic extremes, while individual use areas (second-order) may be highly correlated with the composition and mosaic of distinctive vegetative types available within the geographic range. When spatial areas are considered resource units, the used area at a larger order of selection (e.g., a territory) constrains the scope of availability for the lower order of selection (e.g., resting site). RSAs can be designed to investigate selection at one or several orders of selection, and resource attributes can be characterized at multiple geographic scales. The August 2016

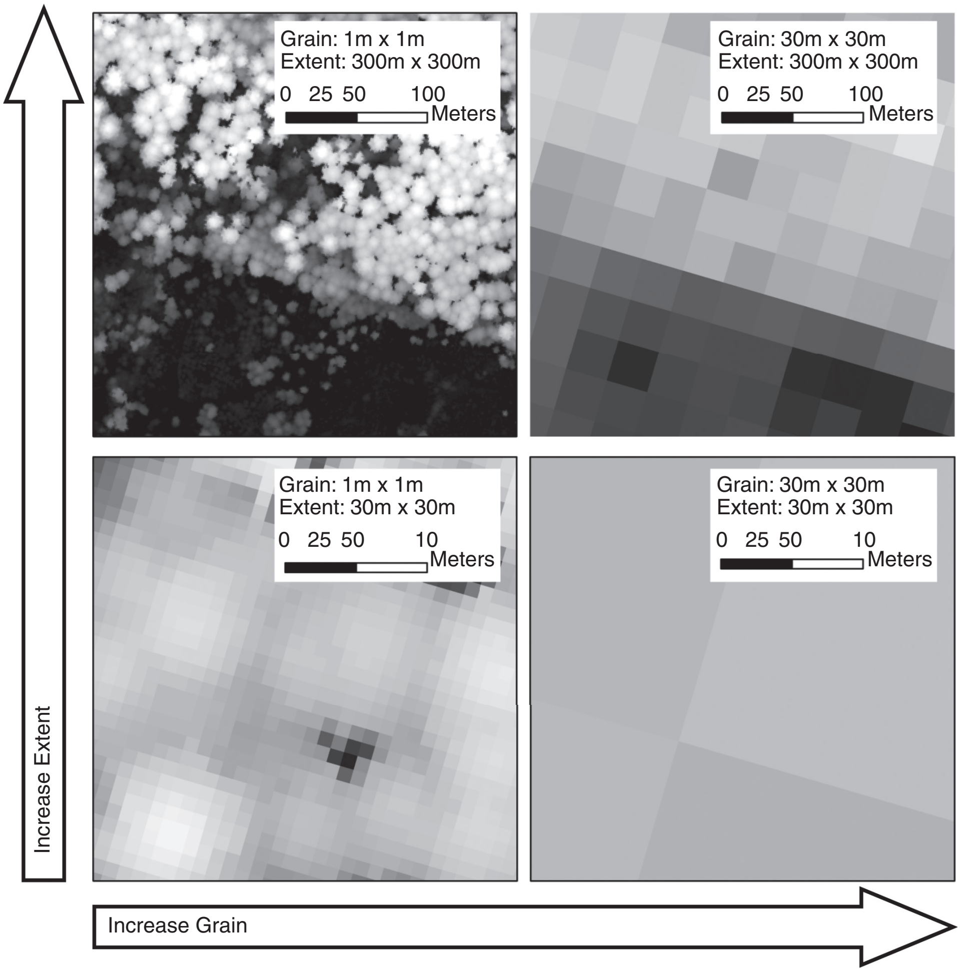


Fig. 12.1. Effect of changing geographical scale of grain and extent in heterogeneous landscapes. Expanding the study extent introduces new resources, and expanding the grain may dilute effects of rare resources.

special issue of *Landscape Ecology* is a good source for further explanations of multi-order and multi-scale resource selection analyses (McGarigal et al. 2016b).

Resources in natural systems are rarely uniformly distributed, and thus studies of resource selection are closely linked to spatial and temporal patterns of the landscape. The scale of a RSA is defined relative to the extent (or domain) and grain (or resolution) of used and available resources (Johnson 1980; Wiens 1989). When resource units represent spatial areas defined by the researcher, the grain chosen should consider the abundance and spatial distribution of use locations in the landscape, as well as the spatial distribution of available resources (Boyce 2006). Interpretations of resource selection can vary widely depending upon the resolution of spatial resource units and what researchers define as being available to animals (Fig. 12.1). For further reading on the importance of scale in RSAs, see Johnson (1980), Wiens (1989), Boyce (2006), Gaillard et al. (2010), and McGarigal et al. (2016a).

Selecting a Resource Selection Analysis

Resource selection analyses represent a diverse collection of statistical methods ranging from simple

Table 12.1. Common resource selection analyses used by wildlife biologists.

Method	Complexity	Attribute data		Study design[a]	Citations
Chi-squared test	Simple	Univariate	categorical	I	Neu et al. (1974)
Rank based tests	Simple	Univariate	categorical	I, II	Friedman (1937); Johnson (1980)
Compositional analysis	Simple to moderate	Multivariate	categorical	II, III	Aebischer et al. (1993)
Logistic regression	Moderate to complex	Multivariate	categorical/continuous	I, II	Manly et al. (2002)
Polytomous logistic regression	Moderate to complex	Multivariate	categorical/continuous	I, II	North and Reynolds (1996)
Poisson regression	Moderate to complex	Multivariate	categorical/continuous	I, II	Manly et al. (2002)
Discrete choice	Moderate to complex	Multivariate	categorical/continuous	III, IV	Cooper and Millspaugh (1999)
Step-selection analysis	Moderate to complex	Multivariate	categorical/continuous	III, IV	Thurfjell et al. (2014)
Occupancy models	Moderate to complex	Multivariate	categorical/continuous	I, II, III	MacKenzie et al. (2002)
MaxEnt	Moderate to complex	Multivariate	categorical/continuous	I, II	Elith et al. (2011)

[a] These categorizations are based on generalities; there are exceptions to every rule, depending upon specific data collection and analysis structure.

metrics for quantifying use to complex multivariate techniques that incorporate spatial and temporal dynamics and variation among groups. Selecting the appropriate analysis can be challenging and should be an integral component of study design. Understanding assumptions inherent to use and availability data collected, methods to collect those data, the objective of the research (e.g., animal space use prediction versus hypothesis testing), and accessibility of methods affect decisions regarding which RSA to use. Here we describe some of the more common or historically influential RSAs in wildlife ecology (Table 12.1), discuss assumptions common in those analyses, and categorize resource selection study designs.

Common Analyses

The chi-square test (Neu et al. 1974) and univariate rank-based methods (Friedman 1937; Johnson 1980) are arguably the most simplistic RSAs in model structure. These analyses compare univariate categorical attribute values to test or rank attribute categories based on null hypothesis of no difference between proportions of used and available units (Alldredge and Ratti 1992; Manly et al. 2002). These methods are used less often today because models have since been developed that rely on fewer assumptions and allow multivariate descriptions of resource units incorporating both continuous and categorical resource attributes.

Compositional analysis of resource selection is common in telemetry studies and other study designs where individual animals are repeatedly observed (Thomas and Taylor 2006). Compositional analysis is an extension of multivariate analysis of variance (MANOVA), which uses individual animals as replicates and can accommodate categorical differences between individuals (e.g., sex, group). The distribution of resources within each animal's home range boundary or other area encompassing relocations is compared to a larger available extent (Aebischer et al. 1993). Compositional analysis is appropriate only when observation sample sizes are large enough for individual use areas to stabilize (Aebischer et al. 1993). This method assumes independence between individuals sampled and normality of covariates (Manly et al. 2002). Results from standard compositional analysis cannot be used to calculate relative probability of use for spatial units (Manly et al. 2002), so weighted compositional analysis refines the method to assign resource use values based on utilization distributions of activity within use areas (Millspaugh et al. 2006). Several reviews have compared compositional analysis with other RSAs (Millspaugh and Marzluff 2001; Manly et al. 2002).

Logistic regression is a type of generalized linear model that relates a linear function of resource attributes to a binary response of used versus available. Logistic regression analyses dominated the resource selection literature in the 1990s and early 2000s. Keating and Cherry (2004) argue that it was often misused. Logistic regression models can incorporate continuous and/or categorical variables, but it is important that they be tested for noncollinearity. Odds ratios are used to interpret influence of attribute coefficients. Standard logistic regression does not account for variable use frequencies; however, polytomous (multinomial) logistic regression and Poisson (log-linear) regression, both generalized linear models, can accommodate variability in the intensity of use frequencies. Polytomous logistic regression relates a nominal (categorical) rather than binary response variable to a linear function of resource attributes (North and Reynolds 1996), and Poisson regression utilizes counts of use (Manly et al. 2002).

Improvements to statistical programs and computing power have led to an increase in hierarchical models (including random effect models) able to support complex model structures accommodating dependencies in use data (e.g., Gillies et al. 2006), multiple orders and scales of selection (McGarigal et al. 2016a), and ecological dynamics (McLoughlin et al. 2010). The incorporation of random effect terms to generalized linear models such as logistic regression models (Gillies et al. 2006) accommodates datasets where samples are not independent or cases where sampling of groups was unequal. Random effects also allow researchers interested in population-level effects to examine variation between individuals (e.g., Thomas et al. 2006).

Conditional availability models are a flexible class of RSAs that can be used whenever available resources differ within a sample of used resources. These models are particularly useful for studies that make repeated observations of marked individuals. Depending upon sampling structure, these models include conditional logistic regression models, matched-case control regressions, discrete choice models, and step-selection models. In discrete choice models, the combination of use location(s) and their matched available units is called a "choice set" (McCracken et al. 1998). These models range from the relatively simple to those that are complex; in the latter random effects in a hierarchical framework can be incorporated in either Bayesian or maximum likelihood approaches (Cooper and Millspaugh 1999; Manly et al. 2002; Thomas et al. 2006; Kneib et al. 2011). Step-selection models can be considered hierarchical extensions of basic conditional availability models in that they combine models of animal movement, which designate available resources for each use location, with integrated models of resource selection (Thurfjell et al. 2014).

Profile models apply statistical distribution measures from observed use locations to infer species use in similar environmental gradients elsewhere. Mahalanobis distance modeling is a profiling technique that uses vectors of average use to assigned values applicable to mapping animal-resource selection based on how similar other areas are to the multivariate mean (Clark et al. 1993). This method can incorporate numerous multivariate attributes and does not depend upon a sample of available habitat (Manly et. al. 2002).

Machine-learning approaches have been adapted to study resource selection as computer systems have improved. These methods are very flexible and can handle highly nonlinear relationships better than more traditional functions such as logistic regression. The maximum entropy method (e.g., MaxEnt) has increasingly been used to evaluate resource selection (Phillips et al. 2017). MaxEnt modeling software uses a machine-learning process to analyze environmental conditions and fit resource selection functions based on observed use contrasted against a large random sample of available resource units (Merow et al. 2013). Many of these more recent analytical methods were developed in conjunction with or for use in GIS. GIS programs are increasingly relevant due to their ability to handle high-volume datasets collected from automated animal and vegetation monitoring systems and for their ability to generate spatially explicit predictions (e.g., habitat maps).

Detection of any wildlife species is imperfect, and for most species detection probabilities will vary among sites. Incorporating detection probabilities in the analysis of site surveys (i.e., probabilistic models of occupancy) provides the most robust and appropriate analytical methodology for population-level studies of resource selection involving nonmarked animals (MacKenzie et al. 2006). It has been repeatedly demonstrated that failing to incorporate imperfect detection can alter forecasted population trends and estimated species distributions, especially for species with low to moderate detection probabilities (e.g., Field et al. 2005; Martin et al. 2005; Rota et al. 2011). There are many site- or time-specific factors that can impede an observer's ability to detect a focal species. For example, sight surveys may have lower detection probabilities during inclement weather if animals are less active or observer visibility is reduced. Further, even if occupancy status is equal among sites, an observer will be less likely to detect the focal species at a site with very dense vegetation compared to a site with greater visibility. MacKenzie et al. (2002) developed a flexible single-season, single-species occupancy modeling framework that has since been built upon to model multi-season (MacKenzie et al. 2003), multi-species co-occurrence (MacKenzie et al. 2004; Rota et al. 2016), abundance (Royle and Nichols 2003), and demographic vital rates (Rossman et al. 2016).

Common Assumptions

All statistical models contain inherent simplifying assumptions to describe complex ecological processes. Two assumptions common to all models of resource selection are: (1) animals display varying degrees of selection for resources at a range of spatial and temporal scales, and (2) use of a resource provides evidence for its importance for the animal's ecology. These assumptions are necessary to infer underlying drivers for selection from observations of use. The choice of additional simplifying assumptions depends upon study designs, study questions, and the method of analysis. Assumptions should be considered when planning or interpreting results in any study of resource selection. Further reading regarding assumptions in RSA using categorical resource classifications can be found in Alldredge and Griswold (2006), assumptions applicable to RSA using telemetry data can be found in Millspaugh and Marzluff (2001), and assumptions inherent in species distribution analyses can be found in Guisan and Thuiller (2005). Some common assumptions used in RSAs include:

1. *Available and used resource units are correctly classified*. This assumption may be violated if there are scale issues or biases in data collection methods. For example, used resources may be misclassified if the location error from a telemetry triangulation is larger than the grain of the categorized resource units. Inherent bias in observations or census data can also be problematic if observed use data is clumped within the available resource area due to nonenvironmental reasons. For example, spatial bias is likely to occur if surveyed areas are not representative of the area considered available to the individual or populations. This potential bias is common in studies that rely on surveys of areas most accessible by observers (e.g., road surveys) and then extend inference to inaccessible areas.
2. *Availability is constant over the period of study*.This assumption may be violated if resource availability changes between years or throughout the seasons, but inference is made more broadly.
3. *Resources or resource units are equally available to individuals within the population*. An examination of species natural history should help determine if this assumption is appropriate and guide data collection of available areas for observed use locations. For example, since dominant individuals

of a territorial species often exclude other individuals from their territory, that area is not available to the entire population.

4. *Selection criteria are constant across individuals.* This assumption is common in studies that do not identify individuals and in studies where repeated samples from individuals are lumped without random effects. This assumption may be violated when selection changes with characteristics of individuals, such as age class, sex, or breeding status.
5. *A random selection of animals was sampled, and those individuals are representative of the population.* This is a basic assumption for most studies interested in population level questions; however, it is not always the case due to the difficulty of sampling populations. This assumption could also be violated if detectability is unequal across sampling areas.
6. *Resource use or selections made by individuals are independent from other individuals.* This assumption may be violated when family members (e.g., mother and offspring) or territorial competitors are included in the same dataset without accounting for dependencies. Some analyses bypass this assumption through the use of random effects on coefficient slopes or by partitioning data.
7. *Relocations of individual animals are not spatially or temporally correlated.* This assumption may be violated when repeated observations are collected in rapid succession.

Study Designs

Resource selection studies can be classified into four general study designs (see Thomas and Taylor 1990; Erickson et al. 2001; Manly et al. 2002). Classification is based primarily on whether resource availability and use are measured at the population or individual level, and whether at least some animals in the population are identifiable. Here we summarize each study design category, list common analyses, and provide an example for each design.

DESIGN I

Individual animals are not identified in design I studies, and available areas are sampled on a population level. This design was common in early RSAs (e.g., Neu et al. 1974) and remains popular for answering questions of selection across large spatial extents. Roadside surveys, such as the North American Breeding Bird Survey (https://www.pwrc.usgs.gov/bbs/), are an example of this study design, but many assumptions are not met so inference is limited. Common analyses include chi-square, logistic regression, log-linear modeling, occupancy modeling, and MaxEnt.

Here we present an analysis that incorporates results from studies by Davis et al. (2016) and Glenn et al. (2017) that used MaxEnt to produce predictive maps (Fig. 12.2) of forests suitable for nesting and roosting by northern spotted owls (*Strix occidentalis caurina*) at two spatial scales. A primary objective of the studies was to generate models to inform regional monitoring and conservation planning, and Glenn et al. (2017) demonstrated an effective method to estimate carrying capacity of dynamic landscapes. Northern spotted owl nesting and roosting locations were collected during long-term demographic research (see Dugger et al. 2016) and land management agency surveys. A quality control process was conducted to ensure use locations were correctly identified and geographically dispersed throughout the entire modeling region. Spatial autocorrelation and sampling bias were addressed by limiting location data to only one location per territorial pair and randomly spacing them no nearer than the estimated median nearest neighbor distance (a.k.a., spatial filtering).

Two orders of selection were investigated. The third-order selection was analyzed at the scale of a forest stand, regardless of patch size or patterns. Available resource unit attributes (predictor variables) were chosen based on forest stand structure and species composition attributes (Davis et al. 2016). An RSF was fit to these data through the use of response functions (e.g., linear, product, qua-

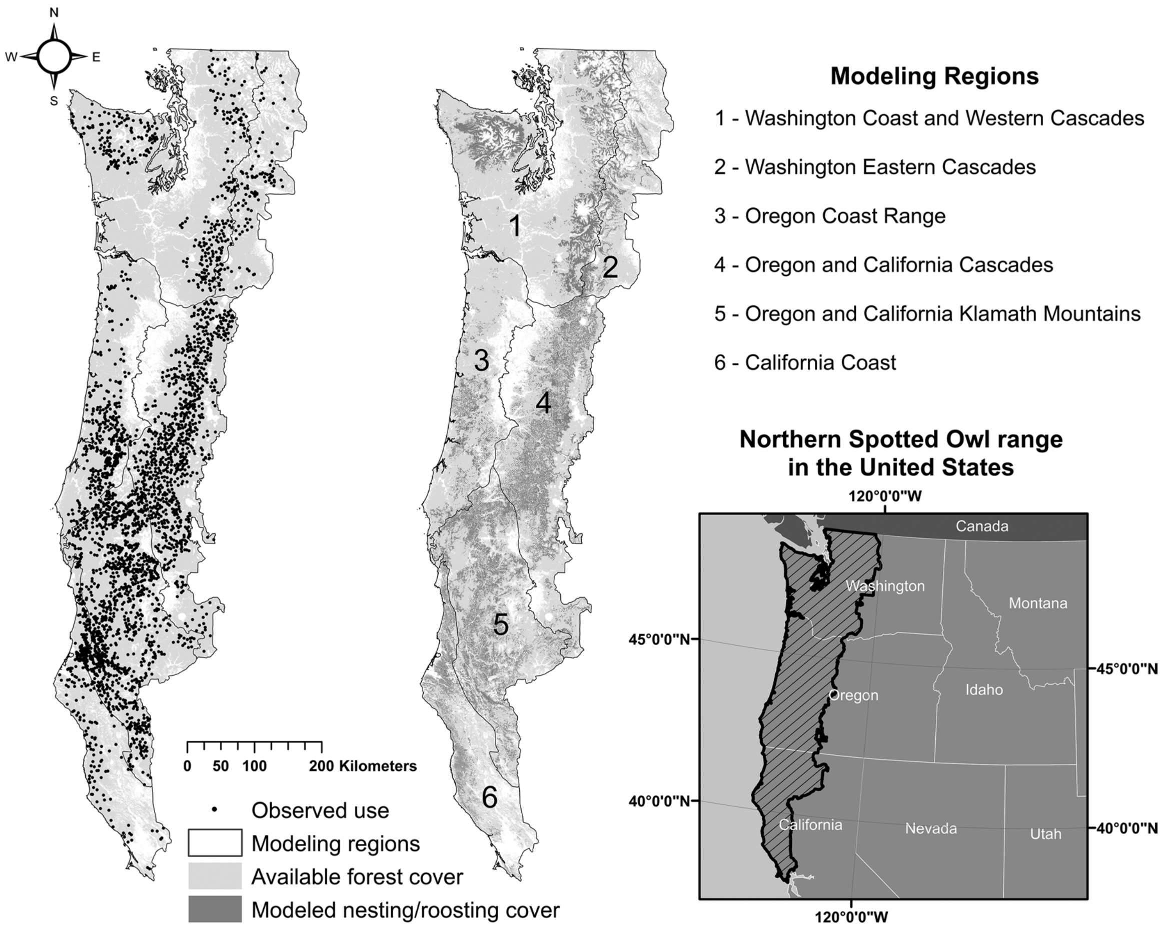


Fig. 12.2. Example of a design I resource selection analysis. Observed use locations and modeling extents used to create northern spotted owl (*Strix occidentalis caurina*) nesting and roosting habitat suitability maps in Glenn et al. (2017) and Davis et al. (2016). *Source: Adapted from Glenn et al. (2017), and used with the permission of Springer Nature.*

dratic) that were determined plausible based on species expert knowledge. Modeling encompassed the entire range of the subspecies, which varied widely in available resources (e.g., redwood forests occurred only in one region of the geographic range). Therefore, the range was subdivided into six modeling regions based on similarity of resources important for northern spotted owls. The models produced represented the relative likelihood of selection of forest types suitable for nesting and roosting use in each modeling region and were mosaicked to produce a range-wide map (Fig. 12.2). At the territory scale (second-order selection) Glenn et al. (2017) used a classified binary version of the Davis et al. (2016) maps to produce predictor variables that represented the amount and spatial arrangement of nesting/roosting forest cover, and also included topographic and climate variables. Forest cover variables were: percentage of forest cover likely to be used for nesting and roosting within various radii (Fig. 12.3a and b), the distance from contiguous large patches of nesting/roosting cover (Fig. 12.3c), and the amount of diffuse (intermixed with edge) nesting/roosting cover (Fig. 12.3d).

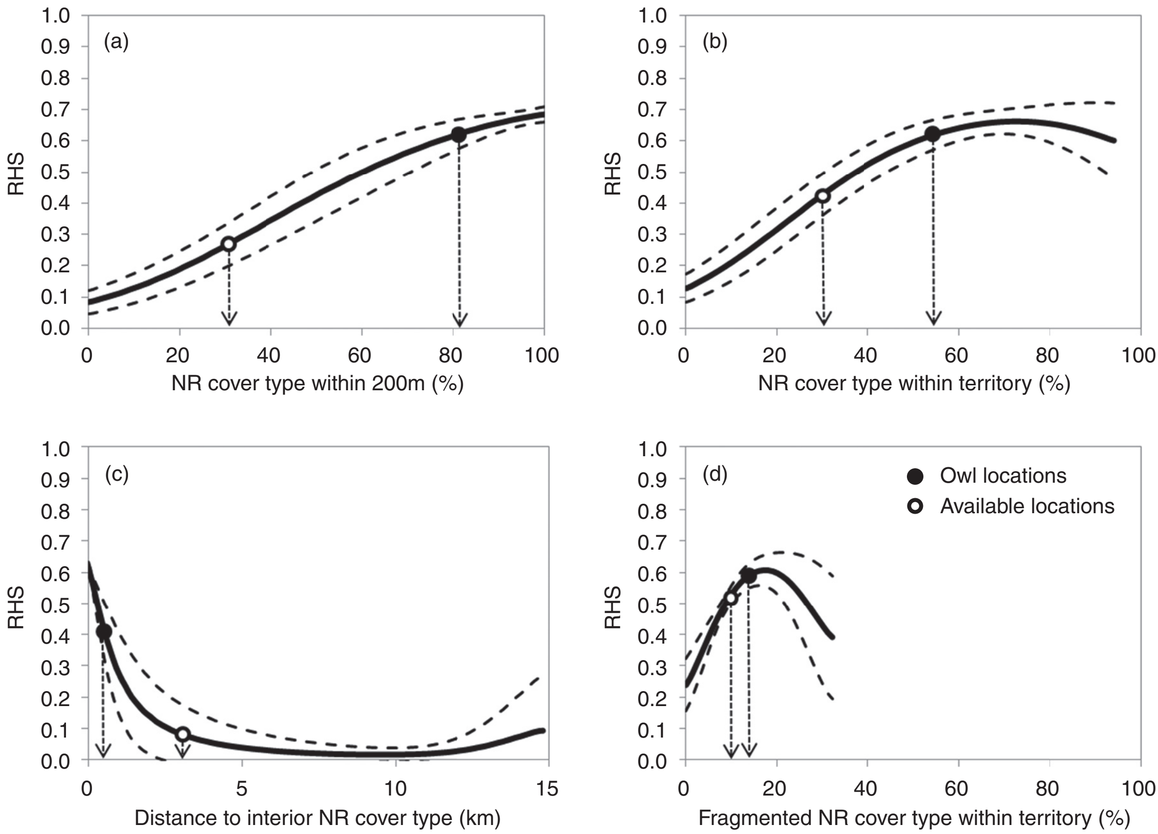


Fig. 12.3. Northern spotted owl (*Strix occidentalis caurina*) relative habitat suitability response functions (solid line) with 95% confidence intervals (dashed lines) based on nesting/roosting cover averaged across modeling regions. The dots represent the average conditions at known spotted owl locations and available locations. *Source: Adapted from Glenn et al. (2017), and used with the permission of Springer Nature.*

The study found selection for landscapes with high cover of nesting/roosting forest types at multiple scales, with a small amount of fragmentation (~20% of the territory) having a positive influence on selection (Fig. 12.3). The same observed use data were utilized for both models, yet the RSFs from each depict the relative likelihood of use at two different spatial scales. It is possible to find forest stands that have all the attributes of heavily used nesting/roosting stands, but have low likelihood of territorial use due their size and juxtaposition in the landscape. Thus, suitability for use at one scale does not always imply suitability for use at another, and in this situation, caution is warranted when representing suitable cover type as suitable habitat.

DESIGN II

In studies with design II, individual animals are identified and resource use is defined for each individual, but resource availability is defined at the population level. This design is common for radio-telemetry studies where relocations are used to describe resource use by tagged individuals, and remotely sensed data are used to describe habitat availability across a study area. Common analyses used with this study design include: Friedman's test, Johnson's method, compositional analysis, discrete choice modeling, logistic regression, log-linear modeling, and multiple regression.

In a telemetry-based study of eastern spotted skunk (*Spilogale putorius*) selection of discrete cover types, Lesmeister et al. (2009) identified resource availability based on proportional coverage of delineated cover types within the study area defined as the maximum convex polygon of all home ranges. The study was conducted in a forested landscape managed primarily for herbaceous understory and an older, open canopy forest. Home ranges were estimated for each individual from 95% fixed-kernel utilization distributions of locations. Selection was determined by a weighted compositional analysis of the proportion of utilization distribution by cover type within each individual's home range (Aebischer et al 1993; Millspaugh et al. 2006). Most available cover types were older pine forest with high herbaceous cover, but eastern spotted skunks showed strong selection for hardwood and young pine stands over other available cover types. In this case, selection for complex forest with dense overstory was likely driven by behaviors to reduce predation risk from avian predators (Lesmeister et al. 2009).

DESIGN III

Studies that identify individuals and quantify resource use and availability for each individual separately are categorized as design III. This study design is common in studies of territorial animals or other situations where monitored individuals would not have equal access to study areas. Common analyses include: compositional analysis, discrete choice modeling, logistic regression, and multiple regression.

Here we present a simplified example of the most common type of discrete choice model used in RSAs, the multinomial logit model, using a subset of northern spotted owl nesting and roosting locations collected on the Klamath, Oregon, demographic study area during long-term demographic research (see Dugger et al. 2016). The primary objective in this example was to determine if fine scale canopy structure attributes generated from remote sensing light detection and ranging (lidar) maps were useful for predicting the selection of a nest site within territories (third-order selection; Johnson 1980). The extent of the analysis was limited to owl territories that overlapped with canopy height map coverages in both space and time. Northern spotted owls are territorial and thus the assumption that areas are equally available for all pairs at the population or regional level is not appropriate; we generated unique choice sets for each territory. The resource units for this analysis were 200-m radius circular plots centered on nest sites (use area) and two random points within the bounds of each territory (sample of available areas; Fig. 12.4).

Discrete choice analysis can accommodate many covariates with variable structure (e.g., linear, quadratic, interactions); however for simplicity, we considered the contribution of just two linear attributes to selection: percentage of mature forest (canopy >80 years) and standard deviation in canopy height. We chose to use a Bayesian framework, similar to that of Jenkins et al. (2017), but these models can also be fit with a maximum likelihood approach (Manly et al. 2002). We used an information theoretic approach, with the Watanabe-Akaike information

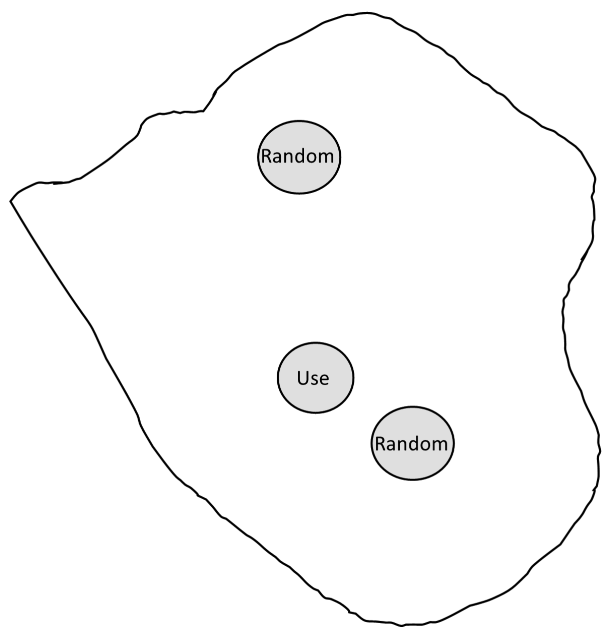


Fig. 12.4. Data collection structure for discrete choice analysis of nest site selection within northern spotted owl (*Strix occidentalis caurina*) breeding territories. Each choice set contained one used location and two randomly selected locations within owl territories.

criterion (see Chapter 4 in this volume), to evaluate single attribute models, a model containing both attributes, and a null model of random selection. The predictive fit of models was calculated using a likelihood ratio test, Estrella's R^2 (Estrella 1998), although a k-fold cross-validation method would have also been appropriate (Boyce et al. 2002). We determined that both covariates were useful in predicting nest locations of northern spotted owls (Table 12.2). Since no interaction effects were present, we can also interpret covariate selection ratios; the selection ratio (exp[coefficient]) measures the multiplicative change in probability of selection when a covariate increases by one unit while others remain the same (McDonald et al. 2006). We found support for positive effects of both percentage of mature forest within 200 m (mean selection ratio = 2.76) and standard deviation in canopy height within 200 m (mean selection ratio = 3.74) on the probability of nest site selection.

DESIGN IV

Studies that make repeated observations of identified individuals, in which resource use and availability are quantified for each observation separately, are categorized as design IV. Resource use is defined for each individual, and availability is uniquely defined for each point of use. Common analyses include discrete choice models and step-selection models.

In the same telemetry-based study of eastern spotted skunk spatial ecology highlighted for study design II, Lesmeister et al. (2008) quantified microhabitat characteristics of sites selected for denning and resting (fourth order of selection). Paired with each used site was a nearby and putatively unused (i.e., available) site suitable for resting or denning. These used and available sites were compared based on habitat characteristics measured at each site. Given the paired used/available design and because reuse of sites by the same study animal was common, Lesmeister et al. (2008) used multinomial discrete choice analysis to fit site selection models in a maximum likelihood framework. In this case discrete choice analysis was most appropriate because the researchers were able to define resource availability separately for each observation (Cooper and Millspaugh 1999). Contrasting used and available sites, they found that selection was based primarily on increased cover of dense vegetation, with additional support for higher rock densities, younger pine forest stands, older hardwood stands, steeper slope, and smaller site entrances. Their results supported the hypothesis that eastern spotted skunks select structurally complex sites for denning and resting partially for thermal regulation, but primarily to enhance protection from predators. This underlying driver in resource selection was also found at the sec-

Table 12.2. Model selection results for discrete choice models evaluating the utility of percent mature forest (mature) and the standard deviation in canopy height (SD canopy) to predict nest area selection by northern spotted owls within territories. Estrella's R^2 is a likelihood ratio test of model fit (1 = perfect prediction, 0 = random).

Model	K[a]	WAIC[b]	ΔWAIC	Estrella's R^2
Mature + SD canopy height	2	85.05	0.00	0.66
SD canopy height	1	96.95	11.90	0.53
Mature	1	112.86	27.81	0.33
Null (random selection)	0	134.04	48.99	0.00

[a] Number of model parameters

[b] Watanabe-Akaike information criterion

ond and third orders of selection (Lesmeister et al. 2009, 2013).

Evaluation and Validation of Resource Selection Analyses

Various methods are used to rank and evaluate the fit of individual models or model sets (e.g., Akaike information criterion (AIC), goodness of fit, etc.). These evaluations are based on the data used to develop (train) the model. However, once the most supported model is decided upon, it is then necessary to validate how well the model predicts use (Fig. 12.5). Using the same observation data to train and test the model may lead to overly optimistic test results and is thus not advised for model validation (Howlin et al. 2003). Using observation data independent of the model training data is considered the best approach. For example, Glenn et al. (2017) used independently collected location data to validate their range-wide model of forests suitable for northern spotted owl nesting and roosting. When independent data are not an option, an often used alternative approach is to subset observation data into model "training" and "testing" replica subsets (e.g., k-fold cross-validation), which can be done in several ways, from geographic partitioning to random selections. A model is then built from each training subset and

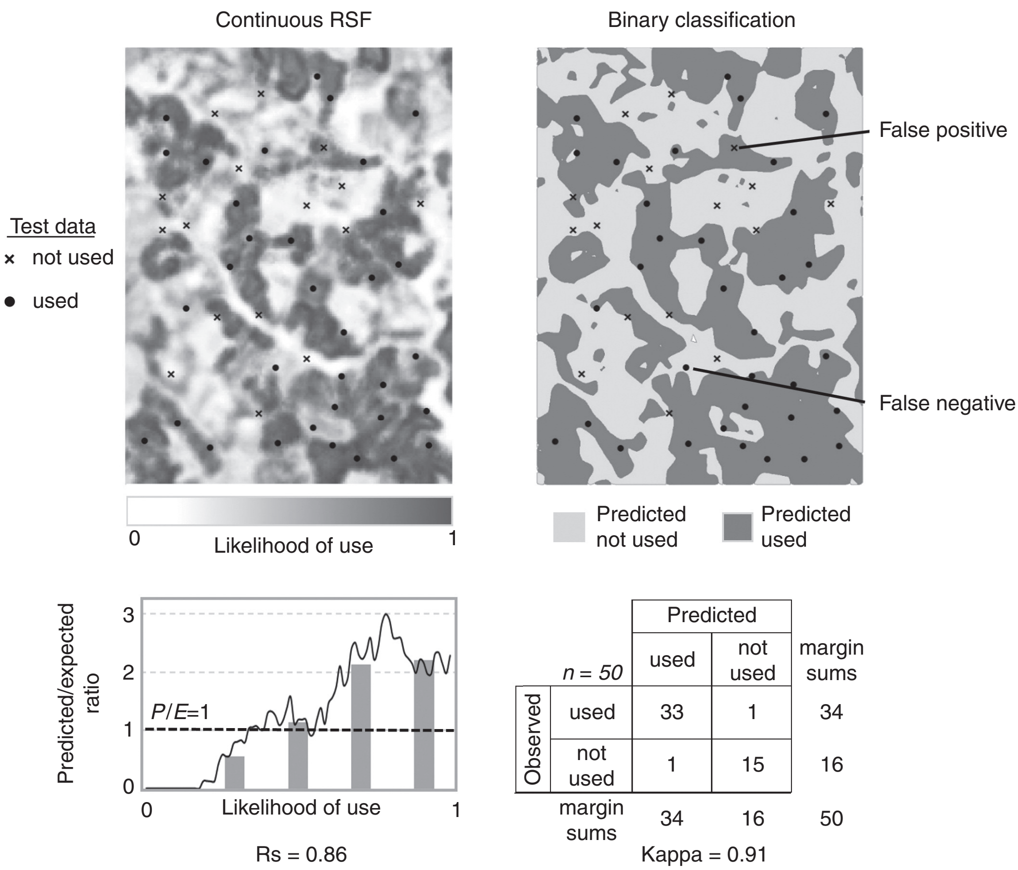


Fig. 12.5. Resource selection analysis map validations showing a continuous resource selection function map (upper-left) and a classified map (upper-right). Beneath each map are examples of proper validation methods, showing predicted-to-expected ratio curve used for continuous RSFs (lower-left) and the confusion matrix used for classified RSFs (lower-right).

its prediction tested with the held-out subset(s). The appropriate model prediction validation method is often dependent on the RSA used. Three commonly used techniques include (1) the confusion matrix, (2) area under the curve (AUC), and (3) area-adjusted frequency (Boyce et al. 2002). Another recent method uses linear regressions between predicted use and observed use, where the slope of the line provides information on how well the model predicted use (Howlin et al. 2003).

Resource selection functions are continuous representations of selection, from low to high. While this gradient of selection is informative, it is sometimes necessary to simplify it into discrete classes for mapping and analysis purposes (e.g., unsuitable or suitable). Where to divide the RSF into these classes is not always clear, and often arbitrary class breaks (e.g., 0.5) are used (Morris et al. 2016). The predicted-to-expected ratio curve method developed by Hirzel et al. (2006) produces information of a model's predictive qualities similar to the area-adjusted frequency method (Boyce et al. 2002), but also provides information on how to classify the model into ecologically meaningful classes in a nonarbitrary fashion (Fig. 12.5). Once an RSA model is divided into discrete classes of use, the confusion matrix can be used for evaluating the model's predictability (Boyce et al. 2002).

Uncertainty is inherent in all models, since they are just representations of observed, complex, and not fully understood processes. Because of this, model validations are a necessary last step in the RSA work flow. Validation methods inform us of the model's usefulness and, just as importantly, weaknesses, for applying the model's predictions of resource selection across space or time.

Summary

The study of resource selection encompasses a vast assortment of modeling techniques and study designs. We have focused our discussion on some fundamental concepts and a few of the most commonly used RSAs from the last few decades to the present. The science of resource selection is rapidly advancing along with improvements in technology and analytical methods, and new methods should be embraced. Hirzel and Le Lay (2008) outlined a dozen useful steps for habitat modelers. A few that we consider most useful to RSA are:

1. Cleary define the goal of the study, particularly in regard to scale and generalization.
2. Preselect resource attributes (variables) considering species ecology and the scale of use data.
3. Carefully delineate the study area to encompass only those resources accessible to the species.
4. Know your chosen analysis method's assumptions, caveats, and strengths.
5. Interpret the model results to ensure they make sense ecologically.
6. Evaluate and assess the model's predictive power and variance.

Regardless of the method, all RSAs are designed to address similar questions and provide useful insights and predictions for wildlife management and conservation. Each method has its strengths and weaknesses, but all are tools that can be used successfully if care is taken in their application.

Acknowledgments

The authors are grateful to the USDA Forest Service Pacific Northwest Research Station and Pacific Northwest Region (Region 6) for supporting their time to contribute to this chapter. They also thank the northern spotted owl demographic team for use of northern spotted owl nest data and E. Glenn for permission to use figures 12.2 & 12.3.

LITERATURE CITED

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

Alldredge, J. R., and J. Griswold. 2006. Design and analysis of resource selection studies for categorical resource variables. Journal of Wildlife Management 70:337–346.

Alldredge, J. R., and J. T. Ratti. 1992. Further comparison of some statistical techniques for analysis of resource selection. Journal of Wildlife Management 56:157–165.

Boyce, M. S. 2006. Scale for resource selection functions. Diversity and Distributions 12:269–276.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling 157:281–300.

Buskirk, S. W., and J. J. Millspaugh. 2006. Metrics for studies of resource selection. Journal of Wildlife Management 70:358–366.

Clark, J. D., J. E. Dunn, and K. G. Smith. 1993. A multivariate model of female black bear habitat use for a geographic information system. Journal of Wildlife Management 57:519–526.

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

Davis, R. J., B. Hollen, J. Hobson, J. E. Gower, and D. Keenum. 2016. Northwest Forest Plan——The first 20 years (1994–2013): Status and trends of northern spotted owl habitats. PNW- GTR-929. USDA Forest Service, Pacific Northwest Research Station, Portland, OR.

Dugger, K. M., E. D. Forsman, A. B. Franklin, R. J. Davis, G. C. White, C. J. Schwarz, K. P. Burnham, J. D. Nichols, J. E. Hines, C. B. Yackulic, P. F. Doherty, L. Bailey, D. A. Clark, S. H. Ackers, L. S. Andrews, B. Augustine, B. L. Biswell, J. Blakesley, P. C. Carlson, M. J. Clement, L. V. Diller, E. M. Glenn, A. Green, S. A. Gremel, D. R. Herter, J. M. Higley, J. Hobson, R. B. Horn, K. P. Huyvaert, C. McCafferty, T. McDonald, K. McDonnell, G. S. Olson, J. A. Reid, J. Rockweit, V. Ruiz, J. Saenz, and S. G. Sovern. 2016. The effects of habitat, climate, and Barred Owls on long-term demography of Northern Spotted Owls. Condor 118:57–116.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43–57.

Erickson, W. P., T. L. McDonald, K. G. Gerow, S. Howlin, and J. W. Kern. 2001. Statistical issues in resource selection studied with radio-marked animals. In Radio tracking and animal populations, edited by J. J. Millspaugh and J. M. Marzluff, pp. 211–242. Academic Press, San Diego, CA.

Estrella, A. 1998. A new measure of fit for equations with dichotomous dependent variables. Journal of Business and Economic Statistics 16:198–205.

Field, S. A., A. J. Tyre, K. H. Thorn, P. J. O'Connor, and H. P. Possingham. 2005. Improving the efficiency of wildlife monitoring by estimating detectability: A case study of foxes (*Vulpes vulpes*) on the Eyre Peninsula, South Australia. Wildlife Research 32:253–258.

Friedman, M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association 32:675–701.

Gaillard, J.-M., M. Hebblewhite, A. Loison, M. Fuller, R. Powell, M. Basille, and B. Van Moorter. 2010. Habitat-performance relationships: Finding the right metric at a given spatial scale. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 365:2255–65.

Gillies, C. S., M. Hebblewhite, S. E. Nielsen, M. A. Krawchuk, C. L. Aldridge, J. L. Frair, D. J. Saher, C. E. Stevens, and C. L. Jerde. 2006. Application of random effects to the study of resource selection by animals. Journal of Animal Ecology 75:887–898.

Glenn, E. M., D. B. Lesmeister, R. J. Davis, B. Hollen, and A. Poopatanapong. 2017. Estimating density of a territorial species in a dynamic landscape. Landscape Ecology 32:563–579.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: Offering more than simple habitat models. Ecology Letters 8:993–1009.

Hirzel, A. H. and G. Le Lay. 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology. 45:1372–1381.

Hirzel, A.H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199:142–152.

Howlin, S., W. Erickson, and R. Nielson. 2003. A validation technique for assessing predictive abilities of resource selection functions. In Proceedings of the First International Conference on Resource Selection, pp. 40–50. Laramie, WY.

Jenkins, J. M. A., F. R. Thompson, and J. Faaborg. 2017. Species-specific variation in nesting and postfledging resource selection for two forest breeding migrant songbirds. PLoS ONE 12:1–14.

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

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

Keating, K. A., and S. Cherry. 2004. Use and interpretation of logistic regression in habitat-selection studies. Journal of Wildlife Management 68:774–789.

Kneib, T., F. Knauer, and H. Küchenhoff. 2011. A general approach to the analysis of habitat selection. Environmental and Ecological Statistics 18:1–25.

Lele, S. R., E. H. Merrill, J. Keim, and M. S. Boyce. 2013. Selection, use, choice and occupancy: Clarifying concepts in resource selection studies. Journal of Animal Ecology 82:1183–1191.

Lesmeister, D. B., R. S. Crowhurst, J. J. Millspaugh, and M. E. Gompper. 2013. Landscape ecology of eastern spotted skunks in habitats restored for red-cockaded woodpeckers. Restoration Ecology 21:267–275.

Lesmeister, D. B., M. E. Gompper, and J. J. Millspaugh. 2008. Summer resting and den site selection by eastern spotted skunks (*Spilogale putorius*) in Arkansas. Journal of Mammalogy 89:1512–1520.

Lesmeister, D. B., M. E. Gompper, and J. J. Millspaugh. 2009. Habitat selection and home range dynamics of Eastern spotted skunks in the Ouachita Mountains, Arkansas, USA. Journal of Wildlife Management 73:18–25.

MacKenzie, D. I., L. L. Bailey, and J. D. Nichols. 2004. Investigating species co-occurrence patterns when species are detected imperfectly. Journal of Animal Ecology 73:546–555.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence. Academic Press, New York.

Manly, B. F. J., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. 2002. Resource selection by animals: Statistical design and analysis for field studies, 2nd ed. Chapman and Hall, London.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. Ecology Letters 8:1235–1246.

Mathewson, H. A., and M. L. Morrison. 2015. The misunderstanding of habitat. In Wildlife habitat conservation: Concepts, challenges and solutions, edited by M. L. Morrison and H. A. Mathewson, pp. 3–8. Johns Hopkins University Press, Baltimore, MD.

McCracken, M. L., B. F. J. Manly, and M. Vander Heyden. 1998. The use of discrete-choice models for evaluating resource selection. Journal of Agricultural, Biological, and Environmental Statistics 3:268–279.

McDonald, L., B. Manly, F. Huettmann, and W. Thogmartin. 2013. Location-only and use-availability data: Analysis methods converge. Journal of Animal Ecology 82:1120–1124.

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

McDonald, T. L. 2013. The point process use-availability or presence-only likelihood and comments on analysis. Journal of Animal Ecology 82:1174–1182.

McGarigal, K., H. Y. Wan, K. A. Zeller, B. C. Timm, and S. A. Cushman. 2016a. Multi-scale habitat selection modeling: A review and outlook. Landscape Ecology 31:1161–1175.

McGarigal, K., K. A. Zeller, and S. A. Cushman. 2016b. Multi-scale habitat selection modeling: Introduction to the special issue. Landscape Ecology 31:1157–1160.

McLoughlin, P. D., D. W. Morris, D. Fortin, E. Vander Wal, and A. L. Contasti. 2010. Considering ecological dynamics in resource selection functions. Journal of Animal Ecology 79:4–12.

Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography 36:1058–1069.

Millspaugh, J. J., and J. M. Marzluff, eds. 2001. Radio tracking and animal populations. Academic Press, San Diego, CA.

Millspaugh, J. J., R. M. Nielson, L. McDonald, J. M. Marzluff, R. A. Gitzen, C. D. Rittenhouse, M. W. Hubbard, and S. L. Sheriff. 2006. Analysis of resource selection using utilization distributions. Journal of Wildlife Management 70:384–395.

Morris, L. R., K. M. Proffitt, and J. K. Blackburn. 2016. Mapping resource selection functions in wildlife studies: Concerns and recommendations. Applied Geography 76:173–183.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 2006. Wildlife-habitat relationships: Concepts and applications, 3rd ed. Island Press, Washington, DC.

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

North, M. P., and J. H. Reynolds. 1996. Microhabitat analysis using radiotelemetry locations and polytomous logistic regression. Journal of Wildlife Management 60:639–653.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017. Opening the black box: An open-source release of MaxEnt. Ecography 40:887–893.

Rossman, S., C. B. Yackulic, S. P. Saunders, J. Reid, R. Davis, and E. F. Zipkin. 2016. Dynamic N-occupancy models: Estimating demographic rates and local

abundance from detection-nondetection data. Ecology 97:3300–3307.
Rota, C. T., M. A. R. Ferreira, R. W. Kays, T. D. Forrester, E. L. Kalies, W. J. McShea, A. W. Parsons, J. J. Millspaugh, and D. Warton. 2016. A multispecies occupancy model for two or more interacting species. Methods in Ecology and Evolution 7:1164–1173.
Rota, C. T., R. J. Fletcher, J. M. Evans, and R. L. Hutto. 2011. Does accounting for imperfect detection improve species distribution models? Ecography 34:659–670.
Rota, C. T., J. J. Millspaugh, D. C. Kesler, C. P. Lehman, M. A. Rumble, and C. M. B. Jachowski. 2013. A re-evaluation of a case-control model with contaminated controls for resource selection studies. Journal of Animal Ecology 82:1165–1173.
Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84:777–790.
Strickland, M. D., and L. L. McDonald. 2006. Introduction to the special section on resource selection. Journal of Wildlife Management 70:321–323.
Thomas, D. L., D. Johnson, and B. Griffith. 2006. A Bayesian random effects discrete-choice model for resource selection: Population-level selection inference. Journal of Wildlife Management 70:404–412.
Thomas, D. L., and E. J. Taylor. 2006. Study designs and tests for comparing resource use and availability II. Journal of Wildlife Management 70:324–336.
Thurfjell, H., S. Ciuti, and M. S. Boyce. 2014. Applications of step-selection functions in ecology and conservation. Movement Ecology 2:4.
Warton, D., and G. Aarts. 2013. Advancing our thinking in presence-only and used available analysis. Journal of Animal Ecology 82:1125–1134.
Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.