A general approach to the analysis of habitat selection

**Thomas Kneib · Felix Knauer ·
Helmut Küchenhoff**

Abstract The investigation of animal habitat selection aims at the detection of selective usage of habitat types and the identification of covariates influencing their selection. The results not only allow for a better understanding of the habitat selection process but are also intended to help improve the conservation of animals. Usually, habitat selection by larger animals is assessed by radio-tracking or visual observation studies, where the chosen habitat is determined for some animals at a set of specific points in time. Hence the resulting data often have the following structure: a categorical variable indicating the habitat type selected by an animal at a specific point in time is repeatedly observed and will be explained by covariates. These may either describe properties of the habitat types currently available and/or properties of the animal. In this paper, we present a general approach to the analysis of such data in a categorical regression setup. The proposed model generalizes and improves upon several of the approaches previously discussed in the literature. In particular, it accounts for changing habitat availability due to the movement of animals within the observation area. It incorporates both habitat- and animal-specific covariates, and includes individual-specific random effects to account for correlations introduced by the repeated measurements on single animals. Furthermore, the assumption that the effects are linear can be dropped by including the effects in nonparametric manner based on a penalized spline approach. The methodology is implemented in a freely available software package. We demonstrate the general applicability and the potential

T. Kneib (✉)
Institute of mathematics, Carl von Ossietzky University Oldenburg, Oldenburg, Germany
e-mail: thomas.kneib@uni-oldenburg.de

F. Knauer
Department of Wildlife Ecology and Management, University of Freiburg, Freiburg, Germany

H. Küchenhoff
Department of Statistics, Ludwig-Maximilians-Universität München, Munich, Germany

of the proposed approach in two case studies: The analysis of a songbird community in South-America and a study on brown bears in Central Europe.

Keywords Categorical regression · Compositional analysis · Habitat selection · Multinomial logit models · Penalized splines · Random effects

1 Introduction

Understanding habitat requirements is one of the most crucial questions in the conservation and management of animals. In many studies, animals are radio-tracked or visually observed and the resulting locations are used to analyze habitat selection (see Manly et al. (2002) for a general introduction and the recent special issue of the Journal of Wildlife Management (Strickland and McDonald 2006) for a review of the current state of the art in habitat selection studies). In this paper, we focus on statistical models that explain the habitat selection process, i.e. describe the habitat type selected by an animal at a given particular time out of a set of options comprising habitat types such as forested areas, open areas and urban areas. To obtain a better understanding of habitat selection, statistical models addressing the following questions are of relevance:

- Is usage selective, i.e. are habitat types used disproportionately to their availability? For example, if forested areas are much more common than open areas in an observation area, does usage of forested areas exceed the larger availability? How can we account for the possibility that some habitat types are not available for specific animals?
- Can individual-specific differences in habitat selection be attributed to covariates describing either characteristics of the animal or of the available habitat types?
- Are the estimation results for this model affected by the repeated measurement design where the same individuals are observed several times, i.e. how can correlations between the measurements of one individual be taken into account?
- How can we differentiate between population-level and individual-level selectivity?
- Is the linearity assumption fulfilled for all covariate effects or do some of the effects require nonparametric modeling?

Based on previous work by Aitchinson (1986), Aebischer et al. (1993) propose Compositional Analysis. Here, proportions of available habitat types are compared to the observed proportions of habitat use (for each animal and for every observed location). This approach allows us to address the first question of differences in habitat selection and avoids problems of correlations by combining all observations of one animal into proportions. This is a suitable strategy if no further information on the animals and the habitat types is available. If, however, influences of covariates are to be analyzed, extended models are required. In particular, if covariates are changing over time, a model based on the proportions is inadequate and statistical modeling should be based on the individual time series observations. In Sect. 4, we consider an example of habitat selection by brown bears where animal-specific covariates such as age or dispersal status and habitat-specific covariates such as exposure to different types of roads are thought to influence habitat selection. The computation of observed

proportions also requires relatively long time series for each animal, and the differing precision caused by unbalanced observation designs has to be taken into account. Moreover, compositional analysis is not applicable when some habitat types are not available for some of the animals.

Logistic regression uses a dichotomized version of habitat selection as the response variable. This is usually presence/absence of a species; in the case of only two habitat types (e.g. forested versus open areas) the binary response can also represent the selection of these habitat types. In the case of presence/absence data, radiolocations are often used as presence data while random points within the home range or the study area (e.g. Poscillio et al. 2004) are used as absence data. There are also examples where the study area is divided into grid cells and presence/absence data is defined based on cells being inside or outside home ranges (e.g. Mladenoff et al. 1995; Schadt et al. 2002). Keating and Cherry (2004) showed how logistic regression could be used in the case of pseudo-absence data. The advantage of logistic regression lies in the possibility of including covariate effects directly, thereby quantifying the influence of covariates on the habitat selection process.

Multinomial logit or discrete choice models, though rarely used for analyzing habitat selection, allow for the simultaneous analysis of several habitat types, and automatically take the dependence in the selection processes into account. Multinomial logit models use only the observed locations as sample units, therefore rendering the debate about the proper generation of absence data in logistic regression obsolete. McCracken et al. (1998) give a detailed introduction to multinomial logit models and discuss a simple example of one female black bear. However, the authors only define availability globally. In their case, availability is only estimated and constant over each year. Arthur et al. (1996) demonstrate with an example of five female polar bears in the Bering and Chukchi seas how to deal with changing availability. They defined availability in terms of circles around each location to account for the highly variable nature of sea ice. Cooper Millspaugh (1999) used multinomial logit models to analyze the selection of 131 day-bed sites by 26 adult elks in South Dakota, USA. The authors used changing availability and different metric and categorical covariates. A recent application of discrete-choice models to northern spotted owl nighttime habitat selection can be found in McDonald et al. (2006).

In all these studies, repeated observations on the same subject are treated as independent, and therefore intra-individual correlations are neglected. In addition, analyses based on repeated measurements bear the risk of attributing individual-level selection to population-level selection. Both problems can be overcome by the introduction of individual-specific random effects. Discrete choice models with random effects are still relatively new even in the statistical literature and therefore receive relatively little attention in applications. A pioneering paper on habitat selection models was published recently by Thomas et al. (2006), who present a Bayesian random effects discrete choice model with inference based on Markov chain Monte Carlo (MCMC) simulation techniques. They model the selection probabilities of grid cells in a discretized observation space in terms of habitat properties such as elevation or land-cover type and allow for individual-specific variation in the corresponding covariate effects. Due to the construction of the discretization, varying availability does not have to be considered, since all grid cells are of the same size. The model does not contain

individual-specific intercepts since these would cancel in the multinomial logit probability transformation.

In our approach, the type of habitat is treated as the categorical response rather than the grid cells within the observation area, in order to gain an understanding of the selection process with respect to the habitat types. In this case, it is important to include individual-specific random intercepts to separate individual-level from population-level selection. In addition, these intercepts should depend on the response category, making them identifiable in a multinomial logit model. Similarly, effects of both animal-specific and habitat-specific covariates will be included in the analyses, which, to the best of our knowledge, have never been considered before and require extended multinomial logit models. Furthermore, it might not always be sufficient to include covariates in purely parametric, linear form in the regression models. We therefore include nonparametric, smooth effects of continuous covariates as a further modeling component in our class of extended multinomial logit models. We present a (penalized) likelihood-based non-Bayesian inferential procedure that might be more familiar to researchers from environmental sciences and avoids questions arising in MCMC-based analyses concerning sensitivity to hyperpriors or mixing and convergence of Markov chains. Note, however, that analogous considerations as in Thomas et al. (2006) could be applied to derive MCMC-based Bayesian estimates in our model, too.

The remainder of the paper is organized as follows: In Sect. 2, we introduce a general non-Bayesian approach for the analysis of habitat selection based on a random effects version of the multinomial logit model. The general applicability and flexibility of the approach are demonstrated in two case studies on 33 individual songbirds from the coastal rain forest in Brazil (Sect. 3) and 22 brown bears from Central Europe (Sect. 4). Section 5 concludes the paper and comments on possible extensions.

2 Modeling habitat selection

We start with the multinomial logit model as a basic discrete choice model for our analysis of habitat selection and assume that there are k different types of habitats. Observations on n animals are collected at different points in time (not necessarily equidistant). The probability that habitat type r is chosen at time t by animal i is denoted by $\pi_{it}^{(r)}$ and is related to the animal- and habitat-specific covariates via

$$\pi_{it}^{(r)} \propto A_{it}^{(r)} \exp\left(\eta_{it}^{(r)}\right) \tag{1}$$

where $\propto$ denotes proportionality up to a multiplicative constant and $\eta_{it}^{(r)}$ is an additive predictor of the form

$$\eta_{it}^{(r)} = \beta^{(r)} + b_i^{(r)} + x_{it}'\gamma^{(r)} + z_{it}^{(r)'}\delta + \sum_{j=1}^{p_1} f_j^{(r)}(u_{itj}) + \sum_{j=1}^{p_2} f_j\left(v_{itj}^{(r)}\right). \tag{2}$$

The components in the model defined by (1) and (2) are given as follows:

$A_{it}^{(r)}$: The availability of habitat type r at time t for animal i. This is a known constant, which is typically proportional to the habitat fraction of the available space. Setting $A_{it}^{(r)} = 0$ for some habitat types r allows us to exclude these habitat types from the choice set. In regression models, $A_{it}^{(r)}$ is called an offset and is introduced to account for the varying availability of the habitat types. After inclusion of the offset, all habitat types virtually are of the same size. The inclusion of availability enables us to analyze whether habitat use is selective, i.e. whether usage is disproportionately compared to availability.

$\beta^{(r)}$ are the parameters of main interest, which indicate the overall habitat preference for the observed animals after accounting for possible covariate effects and availability. Positive parameter estimates indicate an increased use for the corresponding habitat type compared to its availability in the given data. In fact, $\beta^{(r)}$ (as a theoretical quantity in the model) relates to what Johnson (1980) defines as preference: The likelihood of choosing a specific habitat type when it is offered on an equal basis with other habitat types.

$b_i^{(r)}$: To account for correlations between observations on one specific animal i, individual-specific random effects $b_i^{(r)}$ are included. Since all observations on one animal share the same random effect, they are more likely to be alike than two randomly chosen observations, resulting in the desired correlation. The random effects are assumed to be independent and identically Gaussian distributed with category-specific variances, i.e. $b_i^{(r)} \sim N(0, \tau_r^2)$. This corresponds to an equicorrelation model, where all observations on one animal have the same correlation regardless of their distance in time. This may seem quite restrictive in our context but we assume that correlations are mostly due to animal-specific deviations of selection from the overall pattern defined by the fixed effects parameters $\beta^{(r)}$. It is precisely these deviations which are represented by the animal- and category-specific random effects. Defining the expectation of $b_i^{(r)}$ to be zero ensures that $\beta^{(r)}$ can still be interpreted as the overall selection effect since, on average, animals follow this pattern. Assuming the random effects of different animals to be independent implies that correlations are present for observations on the same animal but that observations of different animals remain independent. More complex, time-dependent random effects could be included with the same methodology as employed in this paper if required by the application at hand.

$\gamma^{(r)}$ are the parameters corresponding to the effects of covariates x_{it} depending on time (e.g. daytime of the measurement), and on the animal (e.g. age) but not on the habitat type.

δ are the parameters corresponding to the effects of covariates $z_{it}^{(r)}$, which are dependent on the habitat type r (e.g. distance to a road).

$f_j^{(r)}(u_{itj})$ are nonparametric effects of continuous animal-specific covariates u_{itj} where the functions are category-specific while the covariates are defined

globally. Nonparametric effects are estimated based on penalized splines, see Kneib et al. (2007) and the "Appendix" for details. As is the case with the random intercepts, each function is allowed to have a different amount of smoothness which is estimated from the data together with all other effects.

$f_j\left(v_{itj}^{(r)}\right)$ correspond analogously to nonparametric effects of continuous covariates $v_{itj}^{(r)}$ where the covariate now is category-specific while the effect is defined globally.

The proposed model allows us to address the questions raised in the introductory section. In particular,

- availability is measured and may vary for the subunits (e.g. individuals, locations), see Sect. 3 for possibilities of measuring availability. An offset term is included in the model to account for availability, and it is possible to exclude habitat types from the choice set by setting specific availabilities to zero.
- both categorical and continuous variables can be included as covariates. Covariates may vary over time and are allowed to be either habitat-specific or fixed for all habitat types. In particular, covariates can describe the available habitat types (such as the average distance to roads or elevation, which depend on both the habitat category and the time of the measurement) or the animals (such as sex or age, which are time-constant and time-varying, respectively).
- individual-specific random effects account for intra-individual correlations and allow for differentiation between population-level and individual-level selection.
- effects of continuous animal- and habitat-specific covariates can be included in a non-linear fashion based on penalized splines. The amount of smoothness required for the model is estimated jointly with all other model components, yielding a fully automated smoothing procedure.

Depending on a suitable definition of the covariates, many tools of regression analysis can be employed for the analysis of habitat suitability; we will demonstrate some possibilities later on in the applications.

Since the selection probabilities for the k habitat types have to fulfil the unit-sum constraint (Aebischer et al. 1993), some of the category-specific parameters in the model specification are redundant. We choose the last category as reference and define $\beta^{(k)} = b_i^{(k)} = \gamma^{(k)} = f_j^{(k)}(u_{itj}) = 0$ to ensure identifiability. Furthermore, we set $A_{it}^{(k)} = 1$, i.e. availability is defined relative to the availability of the reference category. To be able to assess the effect of habitat-specific variables $z_{it}^{(r)}$, their values also have to be taken relatively to the value in the reference category. Technically the values $z_{it}^{(r)}$ in the reference category are set to zero and in the other categories $z_{it}^{(r)}$ represents the difference of the variable to the corresponding value for the reference category. Similar restrictions apply for the nonparametric effects of habitat-specific covariates. In this case, the functions $f_j\left(v_{itj}^{(r)}\right)$ have to be centred as $f_j\left(v_{itj}^{(r)}\right) - f_j\left(v_{itj}^{(k)}\right)$. To get a valid model, we furthermore need the assumption that the reference category is available at all points in time t for all animals i, since otherwise we cannot set $A_{it}^{(k)} = 1$.

Once the identifiability restrictions imposed by the unit-sum constraint for the selection probabilities have been accounted for, interpretation of the regression coefficients in the model is most easily accomplished on the level of ratios of probabilities. For the ratio of the probabilities for habitat type r compared to the probability for the reference type, we obtain the expression

$$\frac{\pi_{it}^{(r)}}{\pi_{it}^{(k)}} = A_{it}^{(r)} \exp\left(\eta_{it}^{(r)}\right), \tag{3}$$

i.e. model (1). Accordingly, all covariate effects have to be interpreted multiplicatively on the ratio of probabilities $\pi_{it}^{(r)}/\pi_{it}^{(k)}$ in (3). An alternative interpretation is based on the predictor level, where $\exp\left(\eta_{it}^{(r)}\right)$ may be considered a resource selection function.

Estimation of all components of the multinomial logit model is based on maximum likelihood principles. Inference about model components, e.g. significance tests on some of the regression coefficients, can be based on the large sample properties of these estimates. For our applications, we have conducted a small simulation study to validate that large sample properties can safely be used to draw inferences about the regression coefficients. It is generally hard to draw up rules about the necessary sample size since this depends greatly on design questions such as the number of categories of the response, the distribution of the responses across the categories, or the distribution of the covariates. "Appendix" summarizes inferential procedures and gives additional references.

The approach is implemented in the software package BayesX, which is available free of charge from http://www.stat.uni-muenchen.de/~bayesx. Despite its name, the software also supports the non-Bayesian inference techniques we employ in our approach.

We discuss the relationship of some other models to our approach:

- The model of Arthur et al. (1996) can be seen as special case of our model. There it is assumed that the availability parameters are constant over time for each animal, i.e. $A_{i1}^{(r)} = \cdots = A_{it}^{(r)} = \cdots = A_{iT_i}^{(r)}$. The parameters $\beta^{(r)}$ (ω_k in their notation) are the only parameters in their model. They also use maximum likelihood estimation, which is much easier to handle in the simplified version of our more complex model.

- The methods used by Aitchinson (1986) and Aebischer et al. (1993) are based on a transformation of the data. Our model can be simply rewritten as

$$\ln \pi_{it}^{(r)} - \ln A_{it}^{(r)} = \eta_{it}^{(r)} + const \tag{4}$$

Using the reference habitat k this gives

$$\ln\left(\pi_{it}^{(r)}/\pi_{it}^{(k)}\right) - \ln\left(A_{it}^{(r)}/1\right) = \eta_{it}^{(r)} \tag{5}$$

The corresponding relative frequencies of the left hand side of (5) are modeled by a normal distribution. This approximation has proven useful in practice. The simple model used by Aebischer et al. (1993) is a special case without any covariate effects. Covariate effects could be included, but one would have to assume that all covariates are constant over time, e.g. $x_{i1} = \ldots = x_{iT}$ or $z_{it1}^{(r)} = \ldots = z_{iT}^{(r)}$. Therefore, the approach can be seen as an approximation of our model for the case of fixed further regressors.

3 Case study I: songbirds

3.1 Data collection and processing

The study area of the songbird example is situated in the Mata Atlantica, the Coastal Rain Forest in Brazil. From February 2003 through January 2005, a total of 86 individual Blue Manakins (Chiroxiphia caudata, PIPRIDAE), a small understory omnivorous bird, were captured and radio-tagged. During periods of 10–47 days (depending on battery life of the transmitter), at least one location per individual and day was taken. In our analyses, we only used individuals with at least 10 days of radio-tracking data, i.e. 33 different Blue Manakins. This restriction was required to obtain one habitat type that was common for all observations, which would not have been possible had we included animals with less than 10 observation days. All individuals lived in a fragmented landscape with woodlots of secondary forest and agricultural areas and settlements. We distinguished between the following habitat types: agricultural fields, fallow land, human settlements, young to intermediate forest, old forest and eucalyptus plantation. Old forest was chosen as the reference category (see Hansbauer (2007) for a detailed description).

All data were processed in the Geographical Information System (GIS) ArcView 3.1 (ESRI 1992–1998), and a Mimimum Convex Polygon (MCP) for each individual was calculated using Animal Movement 2.0 (Hooge et al. 1999). This can be considered a home range for resident individuals or the area in which they spent their time during dispersal. Since it is obvious that the available area for each individual is bigger than its MCP, each MCP was buffered with the average distance between consecutive locations, which was 200 m in this example. This area was defined as the available area for each individual. While there are other ways of defining the available areas, e.g. based on Kernel estimates (two-dimensional density estimates), this issue is beyond the scope of this paper. Finally, a database was built with the single locations as records and the percentages of available habitat.

3.2 Models

For the songbirds example, model (1) contains only the habitat preference parameters $\beta^{(r)}$ and the individual-specific random effects $b_i^{(r)}$, yielding the simplified equation

$$\pi_{it}^{(r)} \propto A_{it}^{(r)} \exp\left(\beta^{(r)} + b_i^{(r)}\right). \tag{6}$$

To explore whether the random effect is really required, we also estimated a model without random effects, where $b_i^{(r)} = 0$ is assumed for all animals and all categories, i.e.

$$\pi_{it}^{(r)} \propto A_{it}^{(r)} \exp(\beta^{(r)}). \tag{7}$$

To compare the random effects model (6) and the independence model (7), we consider Akaikes information criterion (AIC) and the generalized cross validation (GCV) statistic (see the "Appendix" for definitions). An alternative model fit criterion would be the log-likelihood. However, the log-likelihood has the same drawback as the R^2 in the linear model, i.e. more complex models always lead to a better fit, reflected by a larger likelihood. To account for model complexity, the equivalent degrees of freedom represent a measure for the effective number of regression coefficients in the model. It is impossible to simply count the additional parameters introduced in the random effects model, since the assumption of normality for the random intercepts effectively reduces the number of free parameters. This is taken into account in a proper definition of the degrees of freedom, see the "Appendix" for details. In the case of the independence model, the degrees of freedom simply reduce to the number of regression coefficients. Both AIC and GCV take the effective degrees of freedom, and thus the increased model complexity in the random effects model, into account. The former is based on a penalized likelihood criterion, where the number of effective parameters is contrasted to the increase in the model fit. The latter is based on an approximation to leave one out cross validation, see the "Appendix" again for a thorough definition. The model fit criteria AIC and GCV allow us to judge whether the increased model complexity does actually count in terms of model fit.

3.3 Results

Table 1 contains the model fit criteria described in the previous subsection for the random effects and the independence model. We find the expected decrease in the likelihood when including random effects but also a considerable increase in the degrees of freedom. For both AIC and GCV there is a substantial decrease, indicating an improvement in the model fit due to the inclusion of random effects. Therefore, we can conclude that random effects are actually required in the model and that individual-specific preferences induce correlations between repeated observations of the same animal.

A more detailed view on individual-specific preferences is possible from the estimated effects $b_i^{(r)}$ displayed in Fig. 1 and the estimated variances of random effects shown in the last column of Table 2. A large value of the estimated variance for a

Table 1 Model fit criteria in the songbirds example

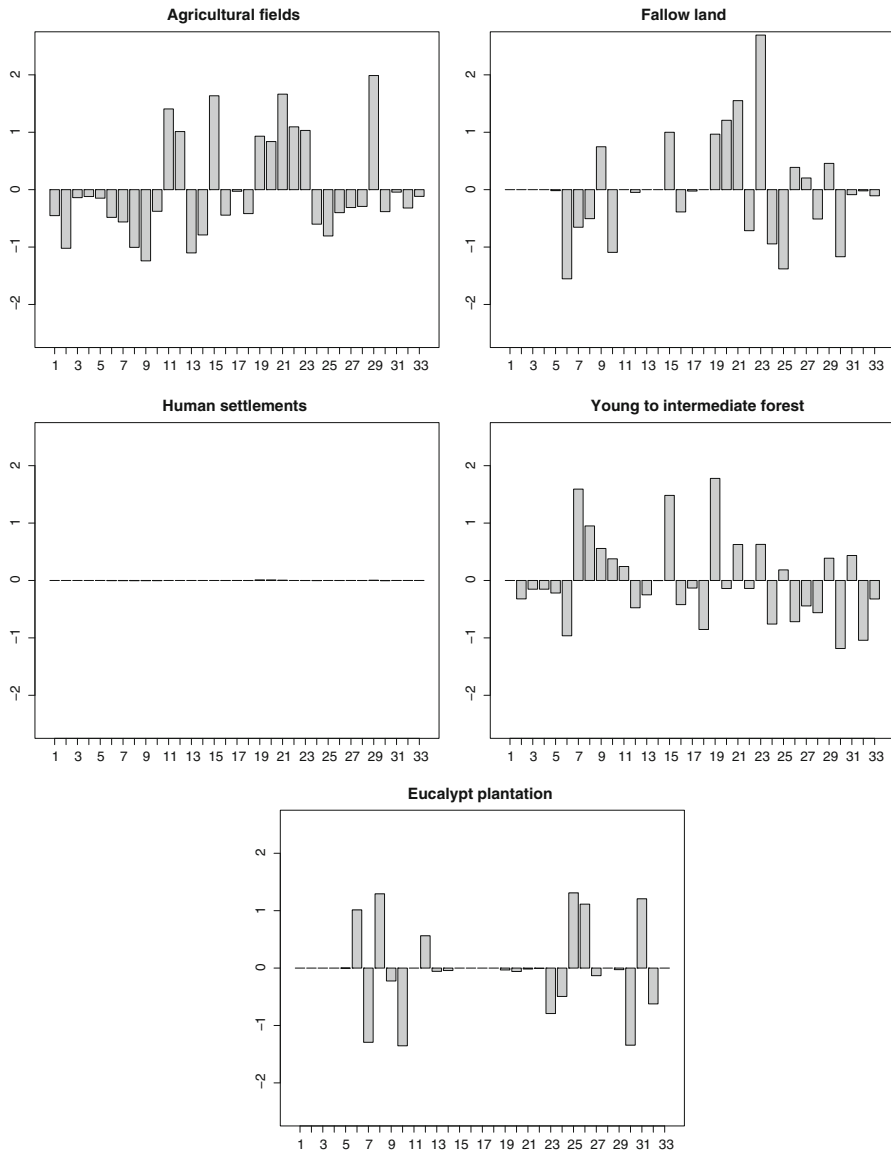|	$-2l$	df	AIC	GCV
Independence model	2205.7	5	2215.7	1.21
Random effects model	1868.5	56.5	1981.5	0.97

Fig. 1 Estimated random effects $b_i^{(r)}$ in the songbird example. The numbers on the x-axis represent the index i of the bird

category r indicates that there is great heterogeneity among the animals in the population for the selection of category r compared to the reference category. Animals with a large positive effect $b_i^{(r)}$ have an elevated preference for this category compared to the "average" bird, whereas animals with negative effects are less attracted by category r. The results in Fig. 1 indicate that heterogeneity is present for all except the human settlement category. While all animals tend to have some individual-specific

Table 2 Estimated preferences $\hat{\beta}^{(r)}$, corresponding standard deviations (SD), 95% confidence intervals, p-values, and variances $\hat{\tau}_r^2$ in the songbirds example

Category	$\hat{\beta}$	SD	95% CI	p-value	$\hat{\tau}_r^2$
Agricultural fields	−3.38	0.361	−4.09 to −2.67	< 0.0001	1.83
Fallow land	−1.38	0.365	−2.09 to −0.66	0.0003	1.81
Human settlements	−4.03	0.411	−4.84 to −3.23	< 0.0001	0.01
Young to intermediate forest	−0.93	0.244	−1.41 to −0.46	0.0003	1.03
Eucalypt plantation	−2.27	0.421	−3.09 to −1.44	< 0.0001	1.64

preferences for agricultural fields and young to intermediate forests, only some specific individuals show up for fallow land and eucalyptus plantation.

The inspection of the estimated preference parameters $\beta^{(r)}$ in Table 2 reveals that the reference habitat type old forest is preferred to all other types. Each of the effects is significant at the 1% level and an overall test for differences between the preference for the categories results in a test statistics of 237.1 on five degrees of freedom, which is also highly significant (see the "Appendix" for details on the construction of the test statistic and the corresponding null distribution). The test statistic in the case of the independence assumption is even larger (610.97) indicating over-optimism when treating the data as independent. In our data example, results in both cases lead to the same conclusion but the additional uncertainty when accounting for correlations in the individuals may well lead to differing results in cases where the p-values are not that clear-cut. Recall also that all results are adjusted for availability of the habitat types by inclusion of the offset terms $A_{it}^{(r)}$.

4 Case study II: brown bears

4.1 Data collection and processing

The study area of the brown bear example is situated in Slovenia, Central Europe, in the Dinaric Mountains. From 1993 to 1998, we captured and radio-tagged 22 different bears. We radio-tracked each individual usually once a day, but dispersing bears in particular were tracked less often due to logistic and financial limitations. In this paper, we only used a maximum of one location per day, in total 1560 locations. The number of observations per bear ranged from only 2 to 198, with most bears (17 out of 22) having at least ten observations and 12 bears having more than 50 observations. The aim of this study was to evaluate the effect of human infrastructure on habitat use by brown bears. Since a pre-analysis showed no difference between open areas and settlements, the two habitat types were combined into one single habitat type (denoted as "open areas" for short). For detailed descriptions, see Kaczensky et al. (2003) and Kaczensky et al. (2006).

For data processing, we proceeded in a similar way to the songbird example. The average distance between measured locations has of course to be recomputed and is given by 2500 m in our data set. In addition to choice and availability information, two types of covariates are available:

- bear-specific discrete covariates were included as four indicator variables: age class with categories yearling, subadult and adult (two indicators), sex (one indicator), and dispersal status with categories resident and dispersing (one indicator).
- habitat-specific continuous covariates are elevation, slope, aspect (circular data, represented as the sum of a sine and a cosine transform in the analyses), and distance to forest roads, paved roads and to the highway, resulting in a total of seven habitat-specific covariates.

For the habitat-specific covariates, average values defined based on the basis of a grid of random points (1 point/ha) were used in the following analyses.

Besides a two habitat model with open areas and forested areas as categories, we considered a second model where forested areas next to roads were defined as additional habitat types. If the area inside forest is within 1000 m of a paved road, then it is classified as "paved road". The areas inside forest and outside "paved road", but within 100 m of forest roads describe the habitat type "forest road". Other areas inside forest and within 1000 m of highways represent "highway". The rest of the forest areas are called "remote forest". We derived the distances to the different road types by comparing the frequencies over the distance of locations with those of random points. In the five-category model we used only the habitat-specific covariates elevation, slope and aspect in addition to the bear-specific covariates as explanatory variables.

4.2 Two habitat types: models

To select the relevant subset of covariates from the set of possible explanatory variables, we performed a backward-forward variable selection within the random effects model

$$\eta_{it}^{(r)} \; = \; \beta^{(r)} \, + \, b_i^{(r)} \, + \, x_{it}' \gamma^{(r)} \, + \, z_{it}^{(r)'} \delta$$

with AIC as optimality criterion. Forested areas were treated as the reference category. It turns out that only distance to forest roads and distance to paved roads appear as influential variables in the final model, see Table 3. For the independence model

$$\eta_{it}^{(r)} \; = \; \beta^{(r)} \, + \, x_{it}' \gamma^{(r)} \, + \, z_{it}^{(r)'} \delta$$

excluding the random effects $b_i^{(r)}$, we did not perform variable selection but re-estimated the model obtained from the random effects model selection procedure. This makes it easier to compare results from both types of models.

Although stepwise model selection procedures have to be used with care and the resulting model should not be accepted blindly as the most suitable model for the data, we consider the stepwise procedure to be useful in our setting for the following reasons:

- The final model obtained from the procedure has been validated with respect to subject matter knowledge in discussions with our collaborators.

Table 3 Estimated preferences $\hat{\beta}^{(r)}$, covariate effects δ, corresponding standard deviations (SD), 95% confidence intervals, p-values, and variance $\hat{\tau}_r^2$ in the brown bears example with two habitat types

	$\hat{\beta}/\hat{\delta}$	SD	95% CI	p-value	$\hat{\tau}_r^2$
Open area	−2.943	0.476	−3.876 to −2.010	< 0.0001	0.40
Distance to forest road	−0.005	0.003	−0.013 to 0.001	0.1166	–
Distance to paved road	−0.001	0.000	−0.001 to 0.000	0.0352	–

Fig. 2 Estimated random effects $b_i^{(r)}$ in the brown bear example with two habitat types
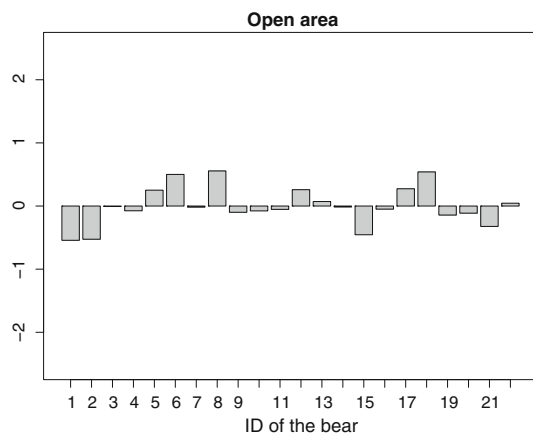
- To obtain estimates for the preferences (which are the parameters of major interest in our model) that are both unbiased and efficient it is not sensible to apply the full model including all covariates. Hence we performed the stepwise selection to find a useful starting model for further consideration. As it is in line with good statistical practice to validate the resulting model, we performed such a validation as noted in the previous point. In the next section, we again take the model delivered from the stepwise model selection procedure as a starting point for further considerations and include nonparametric effects.

4.3 Two habitat types: results

Significant differences between habitat preferences for the two habitat types are identified in both the independence and the random effects model with a forested areas clearly being favored over open areas. The differences between the random effects and the independence model are smaller than for the bird data, as indicated by both the estimated regression coefficients and the magnitude of estimated individual-specific effects (Fig. 2). However, both AIC and GCV still suggest improved model fit by the random effects model.

In order to interpret estimated effects of habitat-specific covariates, we have to relate the difference between the covariate value for open areas and the covariate value in forested areas. For example, for the effect of paved roads, the difference between average distance to paved roads in open areas and average distance to paved roads in

forested areas is considered the covariate of interest and multiplied with the estimated regression coefficient. In our results this coefficient is negative, indicating that a higher distance in open areas compared to forested areas further increases the avoidance of open areas. This unexpected behaviour is, however, supported by bivariate exploratory analyses, where the average distance to paved roads is substantially smaller for bears choosing open areas as compared to bears in forested areas. For forest roads, average distances in both habitat types are comparable in the data set. Both effects are not significant at the 1% level and, as a consequence, no definite conclusion about the effect of distance to roads can be drawn.

4.4 Five habitat types: models

One drawback of the two habitat type analysis with distance to roads as the explanatory variable is that we can only include average distance to roads for the specific habitat type. However, no conclusions can be found about the influence of roads within habitat types. While such effects will be hard to identify from the present data in open areas due to the small number of observations for this habitat type, an extended set of habitat types allows us to address the question in forested areas. We therefore split the original habitat type forested area into the four mutually exclusive areas highway, paved road, forest road and remote forest as described in Sect. 4.1. Remote forest was employed as the reference category and a backward-forward selection procedure was applied to find the best-fitting random effects model

$$\eta_{it}^{(r)} = \beta^{(r)} + b_i^{(r)} + x_{it}' \gamma^{(r)} + z_{it}^{(r)'} \delta$$

according to AIC. Age, elevation and aspect turned out to be important for habitat selection in the five habitat model. For age and aspect, both age dummies and both the sine and the cosine effect were entered into and excluded from the model simultaneously. Based on the best-fitting linear model, we investigated the continuous covariates elevation, sin(aspect) and cos(aspect) for possible nonlinear effects in the model

$$\eta_{it}^{(r)} = \beta^{(r)} + b_i^{(r)} + x_{it}' \gamma^{(r)} + \sum_{j=1}^{p_2} f_j \left(v_{itj}^{(r)} \right). \tag{8}$$

No habitat-specific covariates with linear effects are included, since they are all modeled nonparametrically. On the other hand, all animal-specific covariates are discrete and therefore retain their linear effects.

We found that for all three continuous covariates the effect curves are visually indistinguishable from linear effects, allowing us to focus on the best-fitting linear model in the following discussion. This result should not be misinterpreted as demonstrating the irrelevance of nonparametric models for practical analyses. On the contrary, the more complex model formulation allows us to validate one of the most restrictive assumptions in our model, i.e. the linearity of covariate effects. One of the nice features of
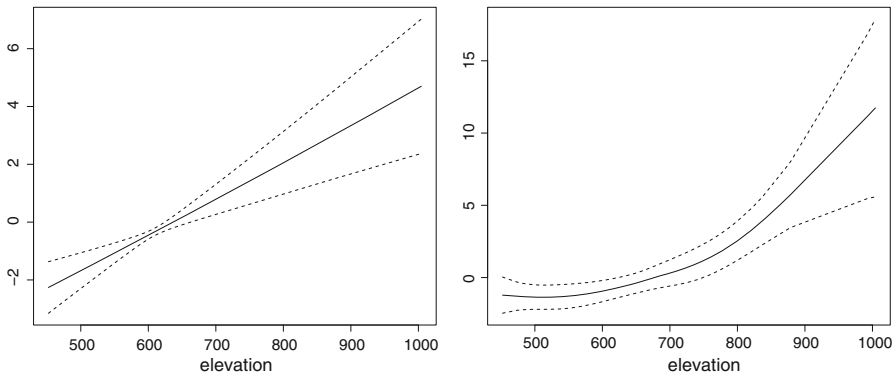
Fig. 3 Estimated nonparametric effect $f(v_{it}^{(r)})$ of elevation (*solid lines*) and 95% pointwise confidence intervals (*dashed lines*) in the brown bear example with five habitat types. The left figure corresponds to the nonparametric effect in the best-fitting model, the right figure shows the nonparametric effects in a full model including all available covariates

the estimation scheme outlined in the "Appendix" is that the resulting estimates indicate when a linear model is sufficient and when more complex, nonlinear modeling is needed. Note also that in a complete model containing all covariates, nonlinear effects of some of the continuous covariates would have been observed. As an example, Fig. 3 shows the nonparametric estimates for elevation obtained in the best-fitting model we consider in the following and in a complete model. While a linear effect is observed in the former case, at least minor deviations from linearity emerge in the latter.

4.5 Five habitat types: results

Tables 5 and 6 summarize the estimated preferences and covariate effects in the best-fitting linear model. For both yearlings and adults, significant differences in habitat preference are found in the data. In contrast, differences for subadults are not significant and the corresponding estimates have relatively large standard deviations. When re-estimating the model without random effects, these standard deviations are dramatically reduced, leading to the false conclusion, that significant differences could be found for subadults also. Note that the point estimates vary to a considerable extent between the random effects and the independence model. This is most clearly observed for the highway preference parameter of subadults, which changes from a large (and significant) positive value in the independence model to a slightly negative (and insignificant) value in the random effects model. Such dramatic changes are not typical in random effects models, where mostly the magnitude of point estimates remains comparable, whereas the standard deviation typically increases. However, in our example, the increase in the standard deviation is so large that the corresponding confidence interval would even contain the value from the independence model.

As in the case of the two habitat analysis, remote forest areas are preferred over open areas. For the forest road habitat type, almost no differences to remote forest are found so that both habitat types are comparable in preference. Results for highways may be

surprising at first sight since our analyses indicate an increased preference compared to purely forested areas for yearlings and adults (i.e. mostly residential bears). However, this habitat type is not only characterized by the presence of noise but also by an almost total absence of human intervention. This seems to attract yearlings and adults, whereas subadults rate the highway habitat type comparable to open areas. The results for paved roads are more difficult to interpret but seem to be caused by collinearity between habitat type and the covariates elevation and aspect. Re-estimating the model without these covariates shows that remote forest is then preferred over all other habitat types and that paved roads have a large negative effect (as expected). Obviously, the distribution of elevation and aspect is not uniform over the habitat types but there is some multicollinearity which leads to the somewhat unintuitive results. This gives a very clear illustration of the importance of carefully checking results for an ecological interpretation, in particular in data sets with a relatively small number of animals such as the one employed in our analyses.

The results in Table 6 indicate a preference for regions with higher elevation. This interpretation is independent of the habitat type considered, since the covariate elevation is category-specific but has a global regression coefficient δ. The aspect effect is difficult to interpret from the estimated coefficients since the circular covariate is modeled as the sum of a sine and a cosine transformation. This results in a sine curve with shift and phase translation, which is visualized in Fig. 4. From this sine curve, we can see that areas with a west-northwest aspect are strongly preferred over areas

Table 4 Model fit in the brown bears example with two habitat types

	$-2l$	df	AIC	GCV
Independence model	335.8	3	341.8	0.21
Random effects model	321.4	8.03	337.4	0.20

Table 5 Estimated preferences, standard deviations (SD) and p-values for the random effects and the independence model in the brown bears example with five habitat types

Age group	Type	Random effects model			Independence model		
		Estimate	SD	p-value	Estimate	SD	p-value
Yearling	Open area	−0.671	0.637	< 0.0001	−0.637	0.501	< 0.0001
Yearling	Highway	1.202	0.686		0.861	0.344	
Yearling	Forest road	0.061	0.204		0.118	0.883	
Yearling	Paved road	1.872	0.595		1.908	0.910	
Subadult	Open area	−1.296	0.607	0.220	−1.101	0.440	< 0.0001
Subadult	Highway	−0.258	0.854		1.667	0.275	
Subadult	Forest road	0.003	0.198		0.086	0.934	
Subadult	Paved road	−0.055	0.551		0.292	0.961	
Adult	Open area	−1.324	0.553	< 0.0001	−1.204	0.295	< 0.0001
Adult	Highway	1.391	0.695		1.907	0.271	
Adult	Forest road	−0.101	0.164		0.010	0.914	
Adult	Paved road	1.352	0.607		1.847	0.917	

The preferences are obtained as the sum of the baseline effects $\beta^{(r)}$ plus the corresponding effects $\gamma^{(r)}$ for the age groups. The p-values correspond to a test on differences in habitat preferences within the age groups as described in the "Appendix"

Table 6 Estimated covariate effects δ, standard deviations (SD), 95% credible intervals and p-values in the brown bears example with five habitat types

	$\hat{\delta}$	SD	95% CI	p-value
Elevation	0.012	0.003	0.007 to 0.018	< 0.0001
sin(aspect)	−0.843	0.206	−1.247 to −0.440	0.0001
cos(aspect)	0.456	0.262	−0.006 to 0.969	0.0808

Fig. 4 Estimated effect of aspect (*solid line*) together with 95% pointwise confidence intervals (*dashed lines*) in the brown bear example with five habitat types
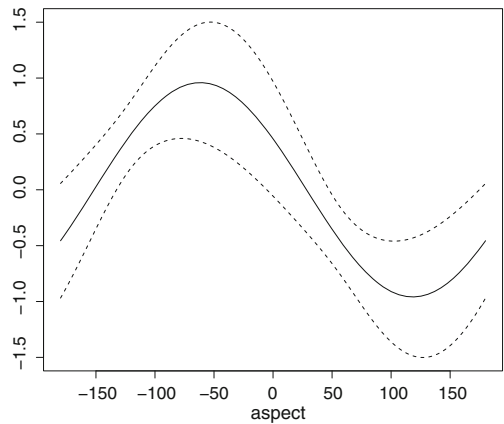

Table 7 Model fit in the brown bears example with five habitat types

	−2l	df	AIC	GCV
Independence model (GLM)	3252.7	15	3282.7	2.05
Random effects model (GLMM)	3128.5	41.5	3211.5	1.90

The first two lines correspond to the best fitting linear model with (GLMM) and without (GLM) random effects. The last two lines correspond to the same model but with nonparametric elevation effects with (GAMM) and without (GAM) random effects

with an easterly aspect (starting with an aspect of zero for north, increasing values correspond to a clockwise increase of the aspect, i.e. 90° corresponds to east, 180° to south, etc.).

When comparing the model fit (Table 7) between independence and random effects model, the largest differences in our three analyses are observed. This again reflects the substantial changes observed in the preference parameters.

Looking at the random effects (Fig. 5), the largest amount of heterogeneity is found for highways. Obviously, some animals tend to prefer areas around highways as habitat type while others tend to avoid these areas. For paved roads and open areas, the individual-specific differences diminish and are no longer present for forest roads.

5 Summary and discussion

In this paper, we presented a general class of categorical regression models which allows us to analyze habitat selection in a realistic modeling framework and to overcome some limitations encountered in previous approaches. Habitat types are treated
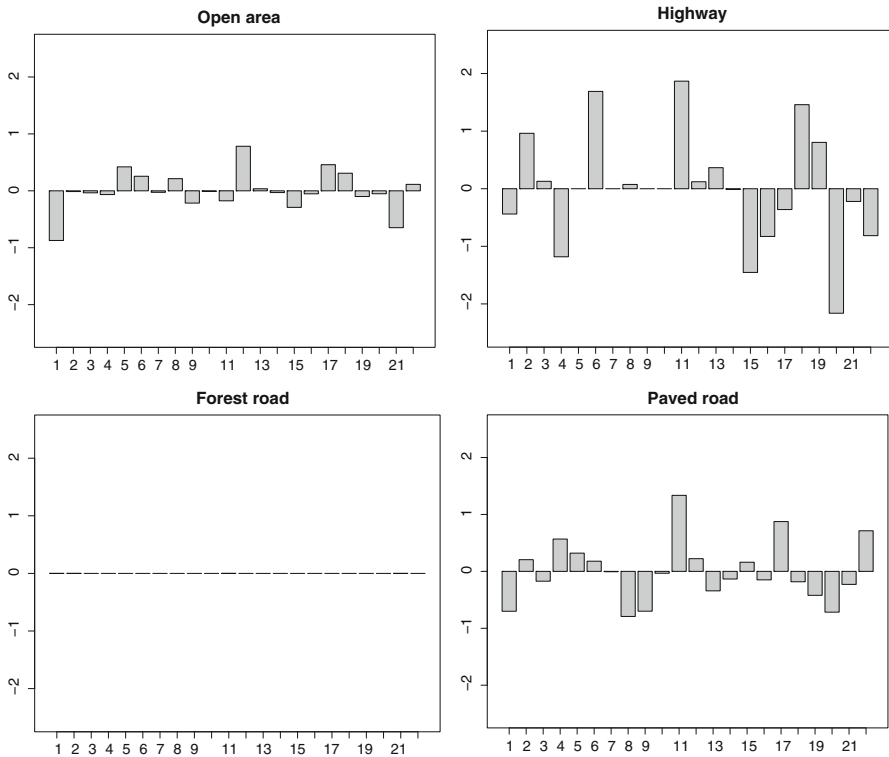
Fig. 5 Estimated random effects $b_i^{(r)}$ in the brown bear example with five habitat types

as a categorical response variable that is related to both animal-specific and habitat-specific covariates. Inclusion of offset terms accounts for varying availability of the habitat types and even for non-availability of some habitat types. Nonparametric modeling of the effects of continuous covariates allows us to validate the linearity assumption underlying the linear multinomial logit model and also has the potential to detect nonlinear functional relationships in larger data sets. To avoid over-optimistic results obtained by treating the repeated measurements of one animal as independent, the model includes random effects for the animals. Random effects also facilitate the separation of individual-level from population-level selection and allow for unbalanced designs where the number of replicated measurements for the animals may vary significantly. This very realistic approach allows us to gain a deeper understanding of the habitat requirements of given species and to evaluate landscapes for these requirements.

Another advantage of the presented approach is that it can easily be extended to more complex situations if needed. For example, we can not only include individual-specific (random) intercepts but also animal-specific covariate effects, e.g.

$$\pi_{it}^{(r)} \propto A_{it}^{(r)} \exp\left(\beta^{(r)} + b_i^{(r)} + x_{it}'\gamma^{(r)} + x_{it}'\gamma_i^{(r)} + z_{it}^{(r)'}\delta\right),$$

where $\gamma_i^{(r)}$ denotes the individual-specific deviation from the population parameters $\gamma^{(r)}$. Similarly, individual-specific departures can be modeled for the parameter δ, or only some of the covariates can be assumed to have individual-specific effects. In our analyses, we did not consider such extensions since only a relatively limited number of animals is available. Extensions that may be considered in future research are the inclusion of a persistence parameter as in Dailey et al. (2007), to take into account the fact that some animals have the tendency to remain in a specific habitat type, or an adjustment for misclassified measurements of the categorical response variable as in MacKenzie (2006). Furthermore, the longitudinal structure of the data could be modeled by including an AR(1) structure for the random animal effect.

One problem of our discrete choice approach is that the choice set consists of few different habitat types and the choice between two regions (e.g. explained by covariates) of the same habitat type is not part of the model. However, this problem can be reduced by extending the number of habitats by including covariates in their definition as demonstrated in Sect. 4.4. We think that the discrete choice approach is particularly useful when the focus is on the habitat choice, i.e. answering the question of habitat type preferences. If one is interested in explaining the movement behavior of animals in detail, then the approach by Johnson et al. (2008) provides a valuable alternative. They model distributions of certain habitat characteristics of the used areas in relationship to the available resources by so-called resource selection functions. These models are typically applied for high resolution data for few animals. Johnson et al. (2008) discuss different approaches, including Ramsey and Usners's persistence model, Gaussian process movement models and discrete choice models in this context.

In summary, the flexibility and extensibility of the proposed modeling framework for habitat selection analyses makes it suitable not only for the presented case studies but also for larger data sets with more complex structures. Embedding the habitat selection process into a statistical model provides us with well-known measures of the model fit and possibilities for model validation.

Appendix: Penalized likelihood inference in multinomial logit models with non-parametric and random effects

The multinomial logit model discussed in Sect. 2 constitutes a multinomial distribution for each of the individual observations $y_{it} = \left(y_{it}^{(1)}, \ldots, y_{it}^{(k)} \right)'$, i.e.

$$y_{it} \sim \mathrm{Mu}(1, \pi_{it}),$$

where $\pi_{it} = \left(\pi_{it}^{(1)}, \ldots, \pi_{it}^{(k)} \right)'$. Hence, the log-likelihood contribution of each of the observations corresponds to the log-density of such a multinomial distribution and is therefore given by

$$l_{it} = \sum_{r=1}^{k} y_{it}^{(r)} \log \left(\pi_{it}^{(r)} \right). \tag{9}$$

Note that in fact the likelihood only depends on the first $q = k - 1$ probabilities due to the unit sum constraint. In our approach, we automatically account for this due to the constraints discussed in Sect. 2.

Under the assumption of independent observations, the joint likelihood would simply be constructed as the sum of all individual contributions, but since we are considering repeated measurements of an individual animal i, the assumption of independence is at least questionable. To account for intra-observational correlations, we introduced individual-specific random effects in (1) and, as a consequence, the likelihood now consists of two parts: the conditional distribution of the responses given the random effects and the random effects distribution. While the former is still of the multinomial form (9) with the random intercept augmented to the model equation as in (1), the latter expresses that the sample of analyzed animals is only a random sample from the overall population. In this paper, we consider Gaussian distributed random effects, i.e.

$$b_i^{(r)} \text{ i.i.d. } N \left(0, \tau_r^2 \right), \quad i = 1, \ldots, n, \tag{10}$$

which is the most common choice in regression models with random effects and reflects the assumption that the factors introducing heterogeneity between the animals are approximately Gaussian distributed in the population. Note that the variances of the random effects depend on the category index r, so that different variability is allowed for the choice probabilities of different categories. In likelihood-based estimation of random effects models, the distributional assumption (10) turns into a penalty term of the form

$$\sum_{i=1}^{n} \frac{1}{\tau_r^2} \left(b_i^{(r)} \right)^2$$

for each of the categories. Hence, (10) acts as a penalty term that penalizes large deviations from the expectation of the random effects distributions, i.e. deviations from zero. As a consequence, maximum likelihood estimates for individual-specific effects can be interpreted as a compromise between fidelity to the data and the prior knowledge expressed through the random effects distribution.

A parsimonious yet flexible approach to estimate the nonparametric effects in (1) can be based on penalized splines (Eilers and Marx 1996; Kneib et al. 2007). Therefore, each nonparametric effect is developed into a series of B-Splines basis functions $B_l(\cdot)$, yielding the expression

$$f(x) = \sum_{l=1}^{L} B_l(x) \zeta_l$$

where ζ_l are the amplitudes associated with the basis functions and both the function and the category indices are dropped for simplicity. To reduce the dependence of the resulting fit on the number of basis functions, a penalty term expressing smoothness assumptions about $f(x)$ is augmented to the log-likelihood. A simple approximation to integrated squared derivative penalties is given by squared difference penalties for the sequence of regression coefficients $\zeta_l, l = 1, \ldots, L$, i.e.

$$\text{pen}(f) = \lambda \sum_{l=d+1}^{L} (\Delta_d(\zeta_l))^2$$

where Δ_d denotes the d-th order difference operator and λ is the smoothing parameter determining the influence of the penalty term. A standard choice for d are second order differences, which are essentially equivalent to a second derivative penalty penalising deviations from a linear effect. Note the similarity between random effects estimation and penalized smoothing: In both cases, a penalty is added to the likelihood to stabilize estimation. While in the case of random affects the penalty is expressed in terms of deviations from the population average, the penalty in the case of a nonparametric function enforces smoothness. In fact, this close connection between random effects and penalized splines also forms the basis for estimating the smoothing parameter jointly with all other parameters. After applying a transformation (see Kneib and Fahrmeir (2006) for details), the penalized spline can also be interpreted as a random effect and the smoothing parameter can be estimated based on mixed model methodology.

To actually compute the maximum likelihood estimates for all regression coefficients in (1), we utilize a Fisher-Scoring algorithm. Let θ denote the vector of all regression coefficients in the model and let $s(\theta)$ and $F(\theta)$ be the first and second derivative of the penalized likelihood with respect to θ. then the Fisher-scoring algorithm proceeds by iteratively updating the current estimates via

$$\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} + (F^{(k)})^{-1} s^{(k)}, \tag{11}$$

beginning with some starting values $\theta^{(0)}$. Upon convergence, the Fisher information matrix $F(\hat{\theta})$ also provides us with the quantities required for the construction of tests and credible intervals for the regression coefficients. Asymptotically, with both the number of observations and the replications per individual large, the ML-estimates are approximately Gaussian distributed and the asymptotic covariance matrix is given by $F^{-1}(\hat{\theta})$. For example, we might test the null-hypothesis of no habitat preference formally via

$$H_0 : \beta^{(1)} = \ldots = \beta^{(k-1)} = 0 \quad \text{vs.} \quad H_1 : \beta_0^{(j)} \neq \beta_0^{(j')} \text{ for some } j \neq j'. \tag{12}$$

Note that the null hypothesis is in fact equivalent to $\beta^{(1)} = \cdots = \beta^{(k)} = 0$ since we initially assumed $\beta^{(k)} = 0$ for identification purposes. The test (12) can be represented

in terms of a general linear hypothesis $H_0 : C\theta = d$ vs. $H_1 : C\theta \neq d$, where C is a full rank matrix. For the particular test (12), we have

$$
C = \begin{pmatrix}
1 & 0 & & \ldots & & 0 \\
0 & 1 & 0 & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & \vdots \\
0 & \ldots & 0 & 1 & 0 & \ldots & 0
\end{pmatrix}
$$

and $d = (0, \ldots, 0)$. For a general linear hypothesis the score test statistic is given by

$$
(C\hat{\theta} - d)'(CF(\hat{\theta})^{-1}C')^{-1}(C\hat{\theta} - d),
$$

and is approximately χ^2-distributed with rank(C) degrees of freedom, i.e. χ^2_{k-1} in our example. In Sects 3 and 4, we used this test to assess the presence of habitat preferences. The last remaining part in the estimation process is the determination of the variance and smoothing parameters collected in the parameter vector Λ. For mixed models, it is common praxis not to estimate these quantities from the joint likelihood of θ and Λ but to use the marginal likelihood for Λ instead, i.e.

$$
L^{\mathrm{marg}}(\Lambda) = \int L(\theta, \Lambda)\mathrm{d}\theta \to \max_{\Lambda}, \tag{13}
$$

where $L(\theta, \Lambda)$ denotes the multinomial likelihood of the model. In Gaussian mixed models, these estimates are equivalent to the well-known restricted maximum likelihood (REML) estimates and could be shown to have smaller bias than ordinary ML-estimates derived from the joint likelihood. In non-Gaussian models, it is not as clear whether marginal likelihood estimates actually perform better than ML-estimates but marginal likelihood estimation also has a nice Bayesian interpretation, which makes them advisable. Proceeding as in marginal likelihood estimation corresponds to an empirical Bayes procedure where the variance components are treated as unknown constant hyperparameters to be estimated from the data. In contrast, the regression coefficients are considered as random variables and appropriate priors are assigned to them. In an empirical Bayes approach, hyperparameters are to be estimated from the marginal predictive density, which (up to proportionality) coincides with the marginal likelihood (13).

Maximization of (13) can again be carried out using a Fisher-scoring type algorithm. First and second derivatives can be derived based on rules for matrix differentiation but we will not discuss this in detail here. A complete description of inferential details in multinomial logit models with additive predictor and random effects can be found in Kneib and Fahrmeir (2006).

Goodness of fit measures for the multinomial logit model can be defined in terms of the deviance residuals

$$
D_{it} = D(y_{it}, \pi_{it}) = 2(l_{it}(y_{it}) - l_{it}(\pi_{it})),
$$

where $l_{it}(\cdot)$ is the log-likelihood of observation i at time t evaluated for either the observation itself or the probabilities π_{it} predicted from the current model. The sum of all deviance residuals is called the deviance

$$D = \sum_{i=1}^{n} \sum_{t=1}^{T_i} D_{it} = 2 \left(\sum_{i=1}^{n} \sum_{t=1}^{T_i} l_{it}(y_{it}) - \sum_{i=1}^{n} \sum_{t=1}^{T_i} l_{it}(\pi_{it}) \right)$$

and based on the deviance and the equivalent degrees of freedom df (see below) we can define the generalized cross validation criterion

$$\text{GCV} = \frac{n}{(n - \text{df})^2} D(y, \hat{\pi})$$

that allows to compare the performance of different models. In Gaussian linear models, the above construction leads to the exact leave one out cross validation statistic, whereas in more general models it can be interpreted as an approximation to this quantity.

The degrees of freedom df associated with a model is given by the trace of the hat matrix projecting the observed responses on their predicted values (see Fahrmeir and Tutz (2001) for details). Again this definition is motivated from Gaussian linear models where this definition results exactly in the number of regression coefficients in the model. In models including random effects and nonparametric functions, the degrees of freedom is a compromise between the number of fixed regression coefficients and the total number of parameters. The exact value is governed by the magnitude of the random effects variances and the smoothing parameters. Based on the degrees of freedom, Akaikes information criterion (AIC) is given by

$$\text{AIC} = -2 \sum_{i=1}^{n} \sum_{t=1}^{T_i} l_{it} + 2\text{df}.$$

and can be used as an alternative measure to compare competing regression models with respect to there model fit.

References

Aebischer NJ, Robertson PA, Kenward RE (1993) Compositional analysis of habitat use from animal radio-tracking data. Ecology 74:1313–1325

Aitchinson J (1986) The statistical analysis of compositional data. Chapman & Hall, London

Alldredge JR, Thomas DL, McDonald LL (1998) Survey and comparison of methods for study of Resource Selection. J Agric Biol Environ Stat 3:237–253

Arthur SM, Manly BFJ, McDonald LL, Garner GW (1996) Assessing habitat selection when availability changes. Ecology 77:215–227

Cooper AB, Millspaugh JJ (1999) The application of discrete choice models to wildlife resource selection studies. Ecology 80:566–575

Eilers PHC, Marx BD (1996) Flexible smoothing using B-splines and penalties. Stat Sci 11:89–121

Fahrmeir L, Tutz G (2001) Multivariate statistical modelling based on generalized linear models. Springer, New York

Dailey M, Gitelman AI, Ramsey FL, Starcevich S (2007) Habitat selection models to account for seasonal persistence in radio telemetry data. Environ Ecol Stat 14:55–68

Hansbauer MM (2007) Responses of forest understory passerines to fragmented landscapes in the Atlantic Rainforest, South-Eastern Brazil. PhD Thesis, University of Freiburg

Hooge PN, Eichenlaub W, Solomon E (1999) The animal movement program. USGS,Alaska Biological Science Center, Anchorage

Johnson DH (1980) The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65–71

Johnson DS, Thomas DL, VerHoef JM, Christ A (2008) A general framework for the analysis of animal resource selection from telemetry data. Biometrics 64:968–976

Kaczensky P, Knauer F, Krze B, Jonozovic M, Adamic M, Gossow H (2003) The impact of high speed, high volume traffic axes on brown bears in Slovenia. Biol Conserv 111:191–204

Kaczensky P, Huber D, Knauer F, Roth H, Wagner A, Kusak J (2006) Activity patterns of brown bears in Slovenia and Croatia. J Zool 269:474–485

Keating KA, Cherry S (2004) Use and interpretation of logistic regression in habitat-selection studies. J Wildl Manag 68:774–789

Kneib T, Fahrmeir L (2006) Structured additive regression for categorical space-time data: a mixed model approach. Biometrics 62:109–118

Kneib T, Baumgartner B, Steiner WJ (2007) Semiparametric multinomial logit models for analysing consumer choice behaviour. AStA Adv Stat Anal 91:225–244

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

Manly BFJ, McDonald LL, Thomas DL, Erickson WP (2002) Resource selection by animals. Kluwer, Doodrecht

McCracken ML, Manly BFJ, Vander Heyden M (1998) The use of discrete-choice models for evaluating resource selection. J Agric Biol Environ Stat 3:268–279

McDonald TL, Manly BFJ, Nielson RM, Diller LV (2006) Discrete-choice modeling in wildlife studies exemplified by northern spotted owl nighttime habitat selection. J Wildl Manag 70:375–383

Mladenoff DJ, Sickley TA, Haight RG, Wydeven AP (1995) Regional landscape analysis and prediction of favorable gray wolf habitat in the Northern Great Lakes region. Conserv Biol 9:279–294

Posillico M, Meriggi A, Pagnin E, Lovari S, Russo L (2004) A habitat model for brown bear conservation and land use planning in the central Apennines. Biol Conserv 118:141–150

Schadt S, Revilla E, Wiegand T, Knauer F, Kaczensky P, Breitenmoser U, Bufka L, Cerveny J, Koubek P, Huber T, Stanisa C, Trepl L (2002) Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. J Appl Ecol 39:189–203

Strickland MD, McDonald LL (2006) Introduction to the special section on resource selection. J Wildl Manag 70:321–323

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

Author Biographies

Thomas Kneib is a Professor for Applied Statistics at Carl von Ossietzky University Oldenburg. He received his PhD in Statistics at Ludwig-Maximilians-University Munich and has been visiting Professor at the University of Ulm and at Georg-August-University Göttingen. He is an Associate Editor of Statistical Modelling, Biometrical Journal and Advances in Statistical Analysis. His research interests include flexible semiparametric regression models, Bayesian regularisation, and Boosting techniques.

Felix Knauer is a Research Assistant at the Department for Wildlife Ecology and Management, University of Freiburg, Germany. He did his MSc on brown bears in Italy and his PhD on bears in Slovenia, both at the Technical University of Munich. Currently he is involved in wolf and bear management in Germany and Austria. His research focuses on habitat analysis and population dynamics of large mammal species.

Helmut Küchenhoff is a Professor at the Department of Statistics, LMU Munich, Germany. He is the head of the Statistical Consulting Unit (STABLAB). His main research interest is applied statistical modeling in

different areas, e.g. biology, medicine and economics. He has been a consultant in many projects on environmental epidemiology and wild biology. Furthermore, he is working on measurement error and misclassification in regression.