Habitat suitability model for the common dormouse (*Muscardinus avellanarius*) based on high-resolution climatic, landscape and forest inventory data

TOBIAS E. REINERS, JULIA NÖDING & JORGE A. ENCARNAÇÃO

Abstract

The common dormouse is a rare, hardly detectable species. Therefore its current distributional range and status still remain unclear. As area-wide distribution data is rarely available we used all existing datasets to build a spatially explicit habitat suitability model in the federal state of Hesse. For environmental data, we used bioclimatic, landscape and forest inventory data on a large-scale but with fine resolution (25 × 25 m). By applying robust statistical methods (Logistic Regression, Boosted Regression Trees) highly reliable models could be constructed. Results showed that the distribution of the dormouse on a large spatial scale can only be assessed by rigorously using all available distribution data in combination with high quality environmental data. With the aid of these models we were able to predict suitability for unknown areas resulting in a spatial explicit habitat suitability map.

Keywords: boosted regression trees, logistic regression, structural richness, distribution analysis, spatial explicit prediction

1. Introduction

Habitat suitability models (HSM) are important tools in species conservation (Guisan & Zimmermann 2000). HSMs can constitute a spatially explicit expansion of the classical niche concept (Austin 2007). Due to technical improvements in Geographic Information Systems (GIS), statistical methods, and availability of high quality remote sensing data, researchers are able to analyze species-habitat relationships on large spatial scales. But the ability of HSMs to describe the species-habitat relationship is heavily dependent on prior knowledge.

At both large and small scale it is known that dormouse distribution is affected by climate (Bright & Morris 1996, Greaves et al. 2006). Temperature and precipitation are driving environmental factors influencing, torpor, activity and reproductive performance of the dormouse (Bright et al. 1996, Juskaitis 2005). Its high dependence on climatic conditions is thought to be complemented by a strong link to natural high forest (Capizzi et al. 2002, Hecker et al. 2003). Additionally, structural factors inside forests influence habitat selection. Light availability, understorey and high diversity of shrubs and herbs are important factors in dormouse habitats (Bright & Morris 1990, Berg & Berg 1998, Capizzi et al. 2002, Panchetti et al. 2007, Juškaitis 2008). Furthermore, in fragmented landscapes dormouse presence is related to the size of forests and amount of hedgerows as either corridors or nesting places (Büchner 2008, Mortelliti et al. 2009).

Having all this information in mind the distribution of the dormouse was analyzed in central Germany in the state of Hesse. In a part of this area (approx. one third) the distribution was additionally analyzed using information from forest inventory data.

2. Material and methods

2.1. Species distribution data

The dataset for this study consisted of 257 presence and 257 absence data points in the federal state of Hesse (Fig. 1A). Data originated from a governmental database (Hessen-Forst FENA – forest inventory and nature conservation agency) with 171 presence points, which mainly consist of data from reporting under the EU Habitats Directive (Büchner et al. 2010). These were supplemented by data from a volunteer survey (the 'Great Nut Hunt', NABU – http://hessen.nabu.de/projekte/nussjagd) with 86 presence points and 257 absence points, where volunteers did not find any signs of dormouse (Fig. 1B). To analyse structural factors within forests, we subsequently analyzed 100 forest stands where dormouse were present or absent in the middle of Hesse (Fig. 1C). Forest inventory data was only available within this area.
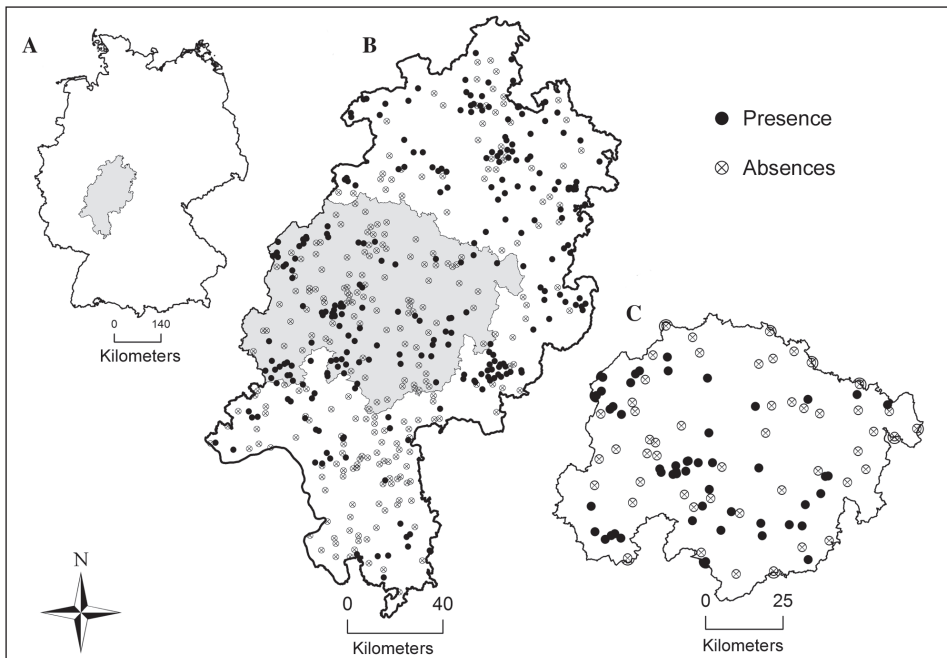
Fig. 1   Spatial scales and distributional datasets used in this study. **A**: Germany with federal state of Hesse highlighted in grey, **B**: Distribution dataset for whole state, **C**: Middle Hesse with presence and absence. Forest inventory data were only available in the Middle Hesse region.

2.2. Landscape data

Selection of environmental data was based on their ecological relevance for the target species (see above). The environmental data used for this analysis comprised: land-use, landscape composition, climate and topography (Tab. 1). For all environmental data gridded raster maps on a resolution of 25×25 m were used. Topography variables included elevation [in m above sea level (a.s.l.)] and slope. The land-use map (ATKIS) consisted of 10 land-use classes (arable land, pastures and grassland, discontinuous urban areas, dense urban areas, deciduous forest, coniferous forest, mixed forests, orchards and coppice, open space, water bodies). ATKIS

stands for Authoritative Topographic Cartographic Information System and is a joint project of all German mapping agencies comprising an object-orientated vector database at the scale of 1 : 25,000. Derived landscape composition variables included percent cover of 8 land-use classes, land-use diversity (Shannon-index) and landscape fragmentation (Interspersion and Juxtaposition Index IJI). Landscape composition variables were estimated within a radius of 1000m and were calculated with SLICER 2.0 (http://www.sfb299.de/slicer).

Five climate variables (BIOCLIM – Bioclimatic variables) were included in the analysis (Hijmans et al. 2005 – http://www.worldclim.org/bioclim). Bioclimatic variables are derived from the monthly temperature and rainfall values in order to generate more biologically meaningful variables. To account for species biology, temperature and precipitation for the warmest period (summer) and for the coldest quarter (winter) as well as temperature seasonality were used.

Additional information about forest stands, each comprising data on tree species, age and percentage of coniferous trees were used for the central part of Hesse. These data were provided by a governmental database (Hessen-Forst FENA).

Tab. 1   Variables used in the state-wide distribution model with data source.

	Variables	Data source
Landscape	10 land-use classes	ATKIS
	8 land-use densities (1000 m)	
	Shannon-Diversity	
	Interspersion-und Juxtapositionsindex	
Topography	Elevation meter a.s.l.	DEM
	Slope	
Climate	Max. temperature (warmest period)	BIOCLIM
	Mean precipitation (warmest period)	
	Mean temperature (coldest quarter)	
	Mean precipitation (coldest quarter)	
	Temperature seasonality	

2.3. Statistical modeling approaches

In this study we used two well-established methods for presence–absence models: Logistic Regression and Boosted Regression Trees (BRTs) (Elith & Graham 2009). In the first instance we used logistic regression (Generalized Linear Model with binomial error structure and logit-link function) starting with a full model including all potentially relevant variables. The number of variables was reduced using AIC (Akaike Information Criterion (Burnham & Anderson 2000) and included afterwards only the set of variables with the most relevant contribution to model fit, leading to the simplest model with highest explanatory value (using stepAIC function in the R statistical program). BRTs were constructed using the library gbm version 1.5–7 in the R statistical program (Ridgeway 2008, Elith et al. 2008, R Development Core Team 2008). As part of the final model, BRT assesses the relative importance (or contribution) of each variable to the final BRT model. This measure is based on the number of times a variable is selected for splitting in the regression trees. As part of the gbm.step each model is evaluated with AUC values with internal cross-validation.

Habitat suitability maps were generated using the ArcMap tool GEPARD 2.1 (Gottschalk et al. 2007; available at www.uni-giessen.de/cms/gepard2) within the ESRI ArcGIS9.3 environment with aid of the spatial analyst extension

In order to assess the relative performance of the models, Area Under the receiver operating characteristic Curve (AUC) were calculated using a 5-fold cross validation and running 100 permutations. AUC values range from 0.5 for models with no predictive ability to 1.0 for models giving perfect predictions.

3. Results

Model results showed that occurrence of common dormouse is related to landscape features and climatic conditions (Tab. 2). The Logistic Regression model had outstanding discriminative ability (AUC = 0.93). Nearly all dormouse presences were found in deciduous forest (n = 67), coniferous forest (n = 32), mixed forest (n = 113) types and in hedgerows (n = 45). Only 10 presences were found outside these land-use classes, mainly in urban areas (towns). It can be considered that this is due to spatial vagueness (25 × 25 m) of land use maps.

All absence points were randomly distributed over land-use classes with the highest number in urban areas (n = 148). Land-use (categorical and continuous variables derived from ATKIS) was the best predictor for dormouse occurrence with relative influence of 79.5 % in BRT models. BRT analyses showed that climate variables significantly explained

Tab. 2 Coefficients of Logistic Regression and BRT relative contributions for the state-wide model. (***P < 0.001, **P < 0.01, *P < 0.05, .P > 0.1).

Variable Group		Coeffizient	Standard error	Significance	Relative Influence in BRT
Intercept	Arable Land	-167.2	87.6	.	
Landscape	Pasture	-0.62	0.93		
	Urban (cities)	-18.12	1.80		
	Urban (towns)	-3.15	0.89	***	
	Hedegrows	2.46	0.73	***	
	Decidious Forest	2.82	0.70	***	79.6 %
	Coniferous forest	3.78	0.86	***	
	Mixed forest	3.69	0.72	***	
	Water	-18.20	2.50		
	residential area in 1km radius [%]	0.03	0.02	.	
	hedgerow area in 1km radius [%]	0.11	0.07		
Climate	Max. temperature in warmest period	2.48	0.91	**	7.6 %
	Max. temperature in warmest period^2	-0.005	0.002	*	
	Precipitation of driest month	-0.87	0.42	*	4.9 %
	Precipitation of driest month^2	0.009	0.004	.	
	Mean temperature of coldest quarter	-0.54	0.17	**	2.5 %
	Temperature seasonality	-0.02	0.01	**	5.4 %

distribution, with highest explanatory value for maximum temperature in the warmest period (7.6 %). Altogether climate variables had a relative influence of 20.4 % in the BRT model. For the state-wide model, Logistic Regression predicted 78.0 % of forest areas to be suitable for dormouse. Differentiation within forests was mainly due to climatic constraints.

The BRT model (AUC = 0.74) based on 50 forest stands with dormouse and 50 forest stands without dormouse revealed that forest age (rel. inf. = 31.8 %), count of tree species (rel. inf. = 17.4 %) and percentage of conifers (rel. inf. = 19.5 %) in combination with maximum temperature in the warmest period (rel. inf. = 31.3 %) were best predictors for habitat suitability. Response curves of BRT analyses showed that forest stands are suitable for the dormouse above an age of 50 years with an optimum at 80 years (Fig. 2A). Furthermore forests are more suitable with more than five tree species present (Fig. 2B) and with a small amount of conifers (up to 9 %, Fig. 2D). The maximum temperature in the warmest period was the only climate variable to be significant on the smaller regional scale with temperatures above 22.5°C promoting high suitability for the dormouse (Fig. 2C).
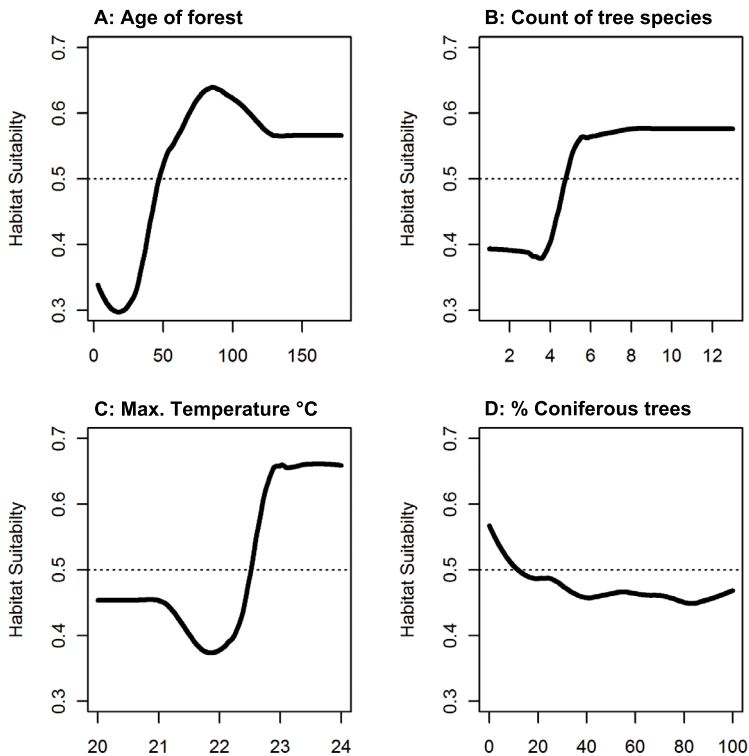

Fig. 2   Partial Response curves of BRT analysis for common dormouse habitat selection for forest stands. Each curve displays the habitat suitability dependent on environmental variables. A habitat suitability value above 0.5 (dashed line) indicates a positive effect for a given variable value.

4. Discussion

Analyses in this study showed that the distribution of the common dormouse is shaped by climatic constraints. Climatic variables were included in a state-wide model as well as regional model. This relationship may be due to torpor behaviour of the dormouse which is triggered by ambient temperature (Bright & Morris 1996, Juškaitis 2005).

State-wide analyses showed that dormouse distribution is not restricted to any specific forest type. Dormice were present in deciduous and coniferous woodland as well as in hedgerows, with some preference for mixed forest. These findings are conform to past studies, where dormice had some preference for broadleaved forest, but were also found in forest with high amounts of coniferous trees (Chanin & Woods 2003, Juskaitis 2007). This study also pointed out that hedgerows are important habitat either as corridors or nesting places. Hedgerows constitute an important habitat element (Ehlers 2009) and landscapes with high amounts of hedgerows are more suitable for the dormouse (Capizzi et al. 2002, Mortelliti et al. 2009). However, state-wide analyses also illustrated that climatic constraints and preference for only four land-use types are not enough to explain the distribution of dormice on a smaller scale.

By using forest inventory data the habitat preference of dormice within forests could be further discriminated. Inside forests, suitability is strongly dependent on tree species diversity, forest age and amount of coniferous trees. These findings were also supported by vegetation mapping inside 12 forest stands, where scrub species richness was highly correlated with tree species richness ($R^2 = 83.4\%$, $p < 0.001$) (Nöding 2011). All three variables promote the idea that optimal dormouse habitats are characterized by high floral
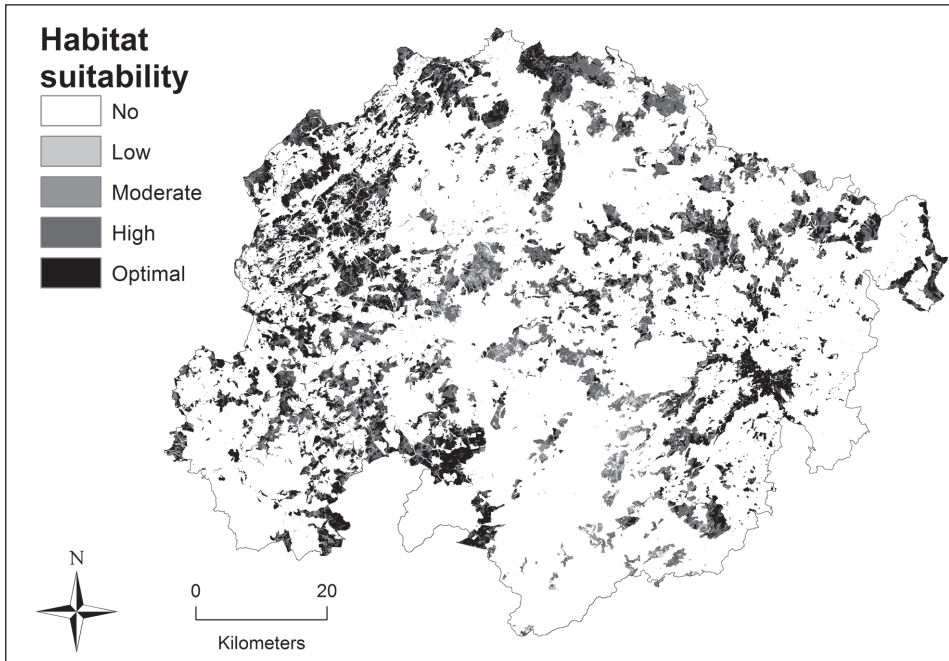
Fig. 3   Suitable forest area in the study region based on forest inventory and climatic data.

richness and a diverse understory. Tree species diversity was highest in forest stands at the medium succession stage (50–100 years). Many tree species together with some coniferous trees supported the development of a diverse understory with many shrub and herb species. In line with other studies, shrub species and cover are the most important factors for optimal common dormouse habitats (Bright & Morris 1990, Berg & Berg 1998, Capizzi et al. 2002). In summary, structural richness inside forest stands was directly associated with tree species composition and forest age. To analyze dormouse habitat distribution at the detailed local level, data on forest stands have to be included.

In addition this study showed that models based on forest inventory data are meaningful but still have limited ability to predict dormouse occurrence and absence (AUC = 0.7). This limitation may be due to additional abiotic and biotic factors influencing dormouse distribution which were not included in models so far (e.g. intraspecific competition, tree hole availability, forest managing practices). These have to be tested in future studies.

This study will provide a basis for future conservation measures and will also be useful in related fields of dormouse research and conservation e.g. the influence of habitat fragmentation. This analysis also highlighted the importance of climatic conditions on a large spatial scale. This will be increasingly important when considering changing climatic conditions.

5. Acknowledgements

6. References

Austin, M. (2007): Species distribution models and ecological theory: A critical assessment and some possible new approaches. – Ecological Modelling **200**: 1–19.

Berg, L. & Berg, A. (1998). Nest site selection by the dormouse *Muscardinus avellanarius* in two different landscapes. – Annales Zoologici Fennici **35**: 115–122.

Bright, P. W. & Morris, P. A. (1990): Habitat Requirements of Dormice *Muscardinus avellanarius* in Relation to Woodland Management in Southwest England. – Biological Conservation. **54**: 307–326.

Bright, P. W. & Morris, P. A. (1996): Why are dormice rare? A case study in conservation biology. – Mammal Review **26**: 157–187.

Bright, P. W., Morris, P. A. & Wiles, N. J. (1996): Effects of weather and season on the summer activity of dormice *Muscardinus avellanarius*. – Journal of Zoology **238**: 521–530.

Büchner, S. (2008): Dispersal of common dormice *Muscardinus avellanarius* in a habitat mosaic. – Acta Theriologica **53**: 259–262.

Büchner, S., Lang, J. & Jokisch, S. (2010): Monitoring the dormouse *Muscardinus avellanarius* in Hesse - a contribution to reporting under the EU Habitats Directive. – Natur und Landschaft **85**: 334–339.

Burnham, K. P. & Anderson, D. R. (2000): Model Selection and Inference. – Springer, 2000.

Capizzi, D., Battistini, M. & Amori, G. (2002): Analysis of the hazel dormouse, *Muscardinus avellanarius*, distribution in a Mediterranean fragmented woodland. – Italian Journal of Zoology **69**: 25–31.

Chanin, P. & Woods, M. (2003): Surveying dormice using nest tubes. Results and experiences from the South West Dormouse Project. – Peterborough.

Ehlers, S. G. (2009): Die Bedeutung der Knick- und Landschaftsstruktur für die Haselmaus (*Muscardinus avellanarius*) in Schleswig-Holstein. – Kiel, Christian-Albrechts-Universität.

Elith, J. & Graham, C. H. (2009): Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. – Ecography **32**: 66–77.

Elith, J., Leathwick, J. R. & Hastie, T. (2008): A working guide to boosted regression trees. – Journal of Animal Ecology 77: 802–813.

Gottschalk, T. K., Ekschmitt, K. & Wolters, V. (2007): GEPARD - Ein GIS-basiertes Modell für die faunistische Beurteilung von Umweltszenarien. – Natur und Landschaft **82**: 306–313.

Greaves, R. K., Sanderson, R. A. & Rushton, S. P. (2006): Predicting species occurrence using information-theoretic approaches and significance testing: An example of dormouse distribution in Cumbria, UK. – Biological Conservation. **130**: 239–250.

Guisan, A. & Zimmermann, N. E. (2000): Predictive habitat distribution models in ecology. – Ecological Modelling. **135**: 147–186.

Hecker, K., Bako, B. & Csorba, G. (2003): Distribution ecology of the Hungarian dormouse species, based on the national biodiversity monitoring system. – Acta Zoologica Academiae Scientiarum Hungaricae **49**: 45–54.

Hijmans, R. J., et al. (2005): Very high resolution interpolated climate surfaces for global land areas. – International Journal of Climatology **25**: 1965–1978.

Juškaitis, R. (2005): Daily torpor in free-ranging common dormice (*Muscardinus avellanarius*) in Lithuania. – Mammalian Biology **70**: 242–249.

Juškaitis, R. (2007): Peculiarities of habitats of the common dormouse, *Muscardinus avellanarius*, within its distributional range and in Lithuania: a review. – Folia Zoologica **56**: 337–348.

Juškaitis, R. (2008): Habitat requirements of the common dormouse (*Muscardinus avellanarius*) and the fat dormouse (*Glis glis*) in mature mixed forest in Lithuania. – Ekológia (Bratislava) **27**: 143–151.

Mortelliti, A., Sanzo, G. S. & Boitani, L. (2009): Species' surrogacy for conservation planning: caveats from comparing the response of three arboreal rodents to habitat loss and fragmentation. – Biodiversity and Conservation **18**: 1131–1145.

Nöding, J. (2011): GIS-gestützte Verbreitungs- und Habitatanalyse für die Haselmaus (*Muscardinus avellanarius*) in Hessen. – MSc-Thesis. – University of Giessen.

Panchetti, F., et al. (2007): Nest site preference of common dormouse (*Muscardinus avellanarius*) in two different habitat types of Central Italy. – Italian Journal of Zoology **74**: 363–369.

R Development Core Team (2008): A language and environment for statistical computing. R Foundation for Statistical Computing, – Vienna, Austria. ISBN 3-900051-07-0 [http://www.R-project.org].

Ridgeway, G. (2008): Generalized Boosted Regression Models. – R package version 1.5-7. [http://CRAN.R-project.org/package=gbm].

Authors' addresses:

Tobias E. Reiners*
Conservation Genetics Group
Senckenberg Research Institute and Natural History Museum Frankfurt
Clamecystrasse 12
63571 Gelnhausen, Germany

Julia Nöding, Jorge A. Encarnação
Mammalian Ecology Group
Department of Animal Ecology and Systematics
Justus-Liebig-University Giessen
Heinrich-Buff-Ring 26
35392 Gießen, Germany

*Corresponding author: Tobias E. Reiners (e-mail: tobias.reiners@senckenberg.de)