Analysis of the regional distribution of epigeic arthropods 3. Distribution models for spiders

Axel Rothländer, Alexander Sührig, Michael Judas und Matthias Schaefer

Synopsis

In an area of continuous beechforests on limestone the distribution of spiders was investigated in a study area of c. 4 km² by pitfall sampling. This paper describes the evaluation of GIS available habitat parameters to develop a predictive model for the distribution of the agelenid spider *Coelotes terrestris*. A two-factor model was derived with the independent variables age of canopy trees and slope differentiated into 3 and 2 classes, respectively. The abundance of *Coelotes terrestris* is predicted in 3 classes of relative activity density, and the model describes 63% of the original 189 trap catches correctly. A factorial sampling design is described that is used to test the model predictions.

Araneae, Coelotes terrestris, modelling, non-parametric ANOVA, sampling design, GIS

1 Introduction

Distribution patterns of epigeic spiders in forests are strongly influenced by structure and quality of the litter layer (JOCQUE 1973, ALDERWEIRELDT & al. 1989). Microhabitat characteristics are determined by the thickness of the litter layer, its microclimate, number and size of hollow spaces, and form and arrangement of litter components (SCHAEFER 1989). These parameters are in general difficult to measure, and it is nearly impossible to have these data available for an entire study area of extended spatial scale. On a large scale, the association of species with certain biotopes allows a macroecological characterization of spiders (HÄNGGI & al. 1995). Degrees of habitat association may be determined by the availability of microhabitats or similarities in general biotic or abiotic parameters.

We adressed the question of predictable distribution patterns of spiders on a meso-scale, i.e. on the landscape level in a c. 4 km² section of continuous beechforests on limestone. For the wider area detailed faunistic information on spiders is available (SÜHRIG 1997). The aim of the project is to analyze correlations of species' occurrences or densities with habitat parameters that are both available for the entire study area, and that are at least indirect indicators of proximate habitat factors. In an extensive survey, ground beetles and spiders were sampled and revealed an array of specific distribution patterns (DORNIEDEN & al. 1996). In the present paper we describe the development of statistical distribution models in the instance of the agelenid *Coelotes terrestris* (WIDER, 1834). From this model a factorial distribution map was produced that is at present object of an evaluation sampling scheme. The corresponding modelling of carabid beetles' distribution is described by JUDAS & al. (1998).

The explicit aim of this study is to model the distribution of *C. terrestris* on a meso-scale. Microhabitat characteristics like micro-scale abiotic conditions and structural ressources, prey availability, or interspecific interactions may influence catch results at a specific trap location. These micro-scale effects interfer with the meso-scale abundance on the sampling plot; they are not incorporated in the models but represent an uncontrolled source of variation in the data.

For the purpose of developing a spatially explicit distributon model for the entire study area it was necessary to use a geographic information system (GIS). This tool allows the combination of different geographic data sources and the presentation of model output in the form of detailed maps. As a potential shortcoming of this research strategy, the explanatory variables are limited to those available in digitized georeferenced databases.

2 Material and methods

2.1 Study area

In the mountainous country in southeastern Niedersachsen a limestone plateau rises east of the Leine valley. This plateau is largely covered by forests, mostly beech forming Carici-Fageta and Melico-Fageta (35% and 65%, respectively, BÖTTCHER & al. 1981). At the southeastern edge of the plateau including its slopes a 3.8 km² study area was selected (cf. Fig. 2) for which detailed data of forestry utilization (»Forsteinrichtung«), soil characteristics (»Standortkartierung«), and topography (from the digital landscape model »DGM5« with 12.5 m resolution) were available. The area rises from the bottom of the slopes at 270 m to c. 420 m a.s.l. on the plateau. General climatic conditions are submontane and subatlantic with an average rainfall of 700 mm per year (DAMMANN 1969) and there is a light continental influence expressed by a yearly range in mean temperatures of 17 K (DIERSCHKE 1989). For a small part of the area soil and vegetation have been studied in detail (cf. THÖLE & MEYER 1979, DIERSCHKE 1989).

2.2 Original sampling

The original survey was carried out with stationary pitfall traps with ethylene glycole (DORNIEDEN & al. 1996). In a factorial sampling design 4 parameters were used to predefine local habitats, namely slope, aspect, soil moisture, and age of canopy trees (DÖRING 1996). Each parameter was devided into 3 classes, and the areas of any combination of these factor classes were determined with the geographic information system ARC/INFO. From 81 possible factor class combinations 69 were detected in the area. Only those areas with a minimum extent of c. 50x50 m^2 and a minimum distance of 25 m to forest edges or forest roads were considered as potential sampling units. For each combination up to 3 replicate areas were chosen for sampling with 1 trap each. As some factor combinations were realized at one or two places only, a total of 189 traps were operated. The analysis is based on the 1-yr-catch from July 1994 through July 1995. During this period traps were exchanged biweekly.

2.3 Species data

A total of >16000 specimens representing 107 species and 18 families were caught during the original survey. 36 species were considered to be adequately sampled by pitfall traps (details in SÜHRIG, unpubl. diploma thesis, University of Göttingen 1996) and comprise 94% of the total catch. 6 species were dominant, namely *Callobius claustrarius* (21% of the 36 species), *Coelotes terrestris* (13%), *Histopona torpida* (9%), *Diplocephalus picinus* (8%), *Coelotes inermis* (8%), and *Pardosa lugubris* (7%). Details of species composition and distribution in the larger area of the »Göttinger Wald« are reviewed by SÜHRIG 1997. In the study area a variety of distribution patterns were revealed, ranging from low to high densities (measured as activity abundance), from low to high frequencies, and from locally restricted to continuous distribution areas (DORNIEDEN & al. 1996).

2.4 Habitat parameters

From the GIS databases describing the study area more parameters were extracted than the 4 used for the stratified sampling design (Table 1). The age of the main canopy layer was derived from forest inventory data, and stands differ in light regime: young and old ones have closed canopies with less light penetrating to the ground compared to medium aged stands. Elevation, slope, and aspect are all derived from a digital topographic landscape model with a resolution of 12.5 m, and all have effects on the local climate. The distance from forest edges was determined from the GIS geometries, and both internal (large forest roads) and external edges are considered as they may be relevant to some species (immigration from surrounding habitats, climatic differences of forest stands close or distant to the edge). Soil moisture classification is derived from the digitized representation of a soil status inventory. This factor was monitored under forest productivity considerations, yet it is of potential relevance to some species' distribution patterns and was of significance in some statistical models (cf. below). Also, it proved to be a useful predictor for the distribution pat-

Table 1

Habitat parameters used in the development of distribution models for epigeic spiders. The parameters available from GIS databases at the original measurement scales were transformed for the use as independent classified variables in distribution models.

parameter	original data		model classification	
	scale	range	levels	classes
age of canopy trees (yr)	numeric	19–169*	3	<30 / 30–120 / >120
elevation (m a.s.l.)	numeric	280–425	2	<380 / >380
distance to forest edge (m)	numeric	25–1191	3	<260 / 260–560 / >560
slope	ordinal	6 levels	2	<5° / >5°
aspect	nominal	8+1** levels	4+1**	N / NE-SE / S / SW-NW
soil moisture	nominal	8 levels	8	unchanged

* age in 1989. ** no aspect if slope = 0° (original data) or if slope < 5° (model).

tern of a carabid species (JUDAS & al. 1998). Nevertheless, it was not further evaluated in this study because of two shortcomings: first, the classification is based on an antiquated scheme which is difficult to reconcile with present forestry classification schemes, and second, some disagreements between these data and the topographic database gave rise to the decision to rely only on the latter for predictive purposes.

The idiosyncratic topography of the study area causes correlations between some habitat factors (Table 2). Steepest slopes are at the lowest elevations ($R_S=-0.66$), there is a tendency of older stands on steeper slopes ($R_S=0.27$), and both low elevations ($R_S=0.41$) and steeper slopes ($R_S=-0.35$) are closer to the forest edges. The original aspect classes differ in their average distance to forest edges, and there are diffences between soil moisture classes with respect to slope, elevation, distance to edges, and aspect classes. These multiple correlations may restrict the possibility to construct general distribution models for transfer to other regions with different sets of factor combinations.

The construction of distribution models is based on the determination of those habitat factors that best describe the realized catch results. For the purpose of developing a spatially explicit representation of a factor model it was necessary to classify the parameters (Table 1) so as to produce sensibly large areas for discrete factor classes. This serves to minimize artefacts that arise from GIS-based combinations of different data layers. Factor classifications are based on the values determined for the 189 trap sites. Age and distance are grouped according to 1/3 quantiles, elevation by the median. Slope is differentiated into ±no and ±strong inclination, aspect is aggregated into the extremes of North and South, and the two intermediate directions of ±West and ±East. Small inclination slopes are classified as ›no aspect‹, and the soil moisture classification was not changed in the analytic steps of this study.

3 Distribution models

3.1 Statistical models

All species data were non-normally distributed, whether on the original scale or after transformation. Therefore, the relevant factors for a descriptive model were extracted by non-parametric methods with the original catch-data as the dependent variable. But the measured activity abundance is subjected to a number of confounding effects and the overall level of population densities is affected by temporal dynamics and geographic differences. Therefore, we consider a relative scale of low, medium, and high abundances as more appropriate for the description or prediction of species' distributions. Consequently, model valuations are based on a contingency table analysis of actual catches and those predicted by the model. For this purpose, the original data were transformed to an ordinal scale of lower, medium, and upper 1/3 quantiles.

In order to reduce the effect of spurious correlations in the dataset the factorial distribution models were restricted to combinations of two independent factors. With the 6 factors available (Table 1), 15 possible factor combinations were tested for effects on the catch results. For each two-factor combination the density values were standardized within the classes of one factor, and a Kruskal-Wallis nonparametric one-way analysis of variance was computed for the standardized values within the classes of the other factor. If this test was significant, the two-factor combination was considered as a potential predictive model. To evaluate the fit of a model with the empirical data, each combination of the factor-classes was attributed to a model abundance class according to the median of catches within the class combination. Thus, the actual catch data classified in 1/3 quantiles can be compared by contingency table analysis with

Table 2
Correlations between habitat parameters at their original measurement scale.

	age	elevation	distance	slope	aspect	moisture
age		-0.03	0.03	**0.27**	*3.0*	*11.9*
elevation	0.67		**0.41**	**-0.66**	*4.3*	***79.3***
distance	0.73	0.0001		**-0.35**	***100.9***	***19.1***
slope	0.0002	0.0001	0.0001		*9.9*	***95.3***
aspect	0.89	0.75	0.0001	0.20		***121.1***
moisture	0.11	0.0001	0.008	0.0001	0.001	

Rank correlations of numeric and ordinal parameters (age through slope), non-parametric 1-way ANOVA of these variables for aspect and moisture classes (χ^2 approximation, df=7), and χ^2 test of association between aspect and moisture classes (df=49). For variable definitions cf. Table 1. Spearman's R_S and χ^2 (in italic) above the diagonal, P values below diagonal. Significant correlations (P<0.01) are in bold.

the predictions of the two-factor model classified correspondingly as low, medium, or high. The results of 9 two-factor combinations that were tested as potential models for the distribution of *Coelotes terrestris* are presented in Table 3. From 9 model combinations only 3 had a contingency coefficient >0.7, and the extreme deviations between data and model, i.e. by more than one abundance class, ranged from 2 to 10.

Table 3
Valuation of primary statistical two-factor models predicting the abundance of *Coelotes terrestris* in 3 density classes. Contingency table analysis results are given for 9 models with interactions of two factors.

factor 1	factor 2	Phi	extreme deviations
age	moisture	0.76	6
age	aspect	0.73	2
age	slope	0.72	4
age	elevation	0.66	5
moisture	aspect	0.66	6
aspect	slope	0.62	7
moisture	distance	0.59	10
elevation	aspect	0.57	10
elevation	distance	0.37	7

Extreme deviations: traps with a difference of 2 abundance classes between actual catches and model prediction.

From the 3 best models in terms of contingency the combination of canopy age and soil moisture was not evaluated any further due to the deficiencies of the soil moisture classification (cf. 2.4 above).

3.2 Corrected factorial model

The above classification of factor combinations produces a statistical model that may not be valid in a biological sense. The two alternative models of age/aspect- and age/slope-combinations had to be checked for logical inconsistencies and some appointments of class combinations to predicted model abundance classes had to be revised. Thus, from the primary statistical model a consistent descriptive model was derived. This process increased the number of false predictions and decreased the degree of contingency. In the case of the age/aspect model the extreme deviations increased from 2 to 15 (Phi=0.54). The fit of the modified age/slope model was only slightly less than the primary statistical model, with Phi=0.71 and the same number of 4 extreme deviations (Fig. 1).

Thus, from a range of 9 potential two-factor models a combination of canopy age classes and slope classes was selected as the best descriptors for the original catch data. 63% of all trap results are described correctly, 34% deviate by 1 abundance class, and 2% by 2 classes (cf. Fig. 1). But 60% of the traps

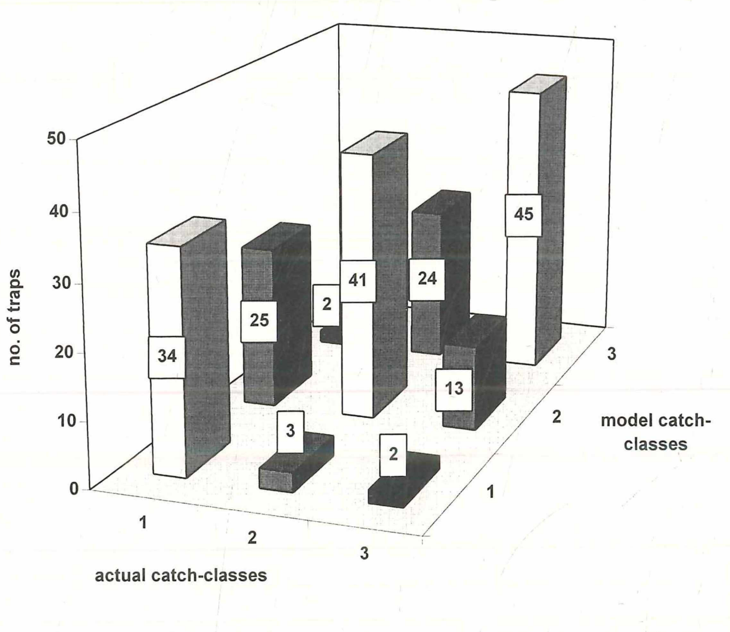

Fig. 1
Agreement of *Coelotes terrestris* actual catches with catches expected from the factorial model. The description by the model is correct (white bars) or deviates from the actual catches by 1 (grey bars) or 2 classes (black bars). $\chi^2=106.8$ (df=4, P=0.001), $\gamma=0.84$ (±0.04 asymptotic SE).

with a mismatch of model prediction and actual catch differ by only 1 or 2 specimens and may thus be regarded to be »nearly correctly« predicted. 2 out of 4 extreme deviations may be explained by insufficient trap efficiency, as less specimens were caught than predicted. The other 2 severe deviations from the model cannot be explained as artefacts because many specimens were caught where the lowest number is expected. Overall, the model predictions can be regarded as a good representation of the quantitative distribution pattern of *Coelotes terrestris* in the study area.

3.3 Model evaluation

From the factorial distribution model a spatially explicit distribution model can be derived that predicts abundances for the entire study area (Fig. 2). There is a large degree of correspondence between the original data and the model (Figs 1 & 2). This is a prerequisite of an acceptable model, yet its validity has to be tested by independent sampling. This is the subject of current sampling, and the design of the evaluation study is to be outlined here.

The combinations of canopy tree age and slope (as defined in Table 1), their prediction of *Coelotes terrstris* relative abundances, as well as their specific correct predictions and strong deviations are given in Table 4. Low abundances are predicted for young or old stands on steep slopes, high abundances for medium aged stands in ±flat areas. The best correspondence of >80% between model and original data is for low densities, the least adequate description is for young stands with little inclination where the false predictions outnumber the correct ones. In order to allow a potential model modification and improvement a stratified sampling design was devised for the model evaluation: 4–12 discontinuous areas of 2500 m^2 minimum size were identified for each factor class combination by GIS data layer intersection, and for each area a number of 1–5 traps according to the size of the area was selected with 25 m minimum distances to forest edges and borders with areas of other factor combinations. This design with a total of 75 traps operated during the main activity season of

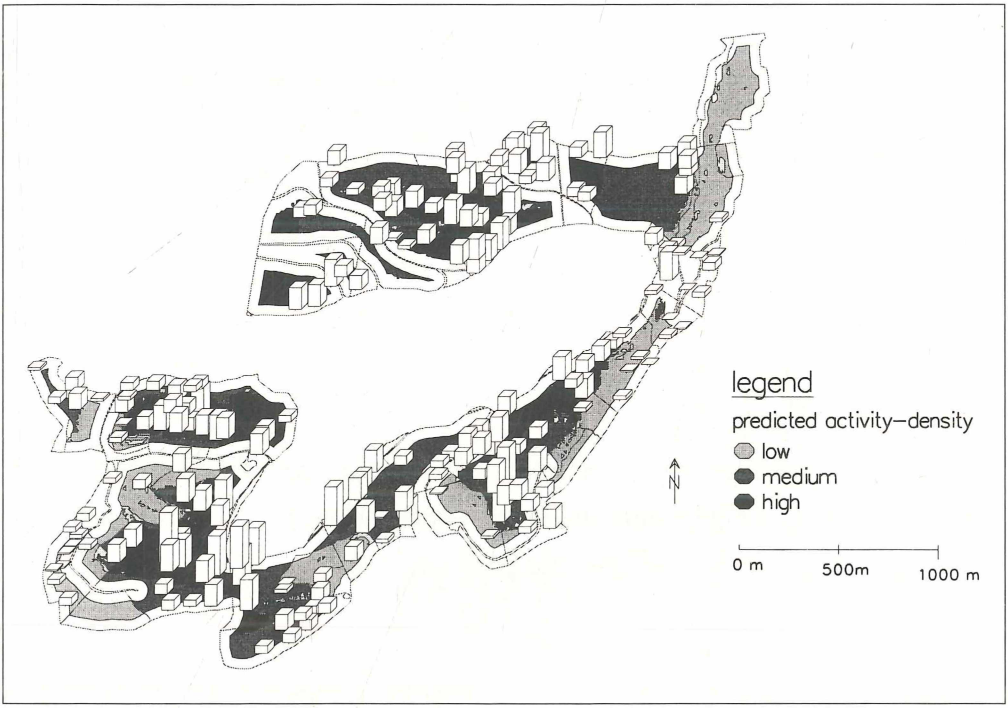

Fig. 2
Study area with the prediction of 3 density classes for the distribution of *Coelotes terrestris*. White columns represent the catch results of the original survey (maximum height corresponds to 39 specimens per trap). White areas are stretches of 25 m width adjacent to borders that were excluded from test sampling. The outlined area is situated at the southeastern edge of the limestone plateau »Göttinger Wald«, and is part of the forestry district of Reinhausen-Wittmarshof.

Table 4

Factorial distribution model of *Coelotes terrestris:* data-model correspondences, distinct areas of different factor combinations, and test sampling trap number allocations to factor combinations. Three abundance classes are predicted from combinations of canopy tree age and slope.

age	slope	predicted abundance	data-model agreement			evaluation design	
			traps	correct	diff. >2	areas	traps
young	flat	medium	12	33%	42%	4	5
	steep	low	14	93%	7%	4	10
medium	flat	high	71	63%	27%	10	25
	steep	medium	58	55%	17%	12	12
old	flat	medium	9	56%	22%	4	5
	steep	low	25	84%	8%	9	18

Age classes are delimited by 30 and 120 yr, slope classes by 5° inclination. diff: >2 specimens difference between actual catch and model prediction.

Coelotes terrestris allows to discriminate between GIS available habitat parameter combinations and their adequate prediction of relative activity abundances.

Acknowledgements

This study is part of a project based at the Forest Ecosystem Research Centre of the University of Göttingen (FZW) and was funded by the Federal Ministry of Education and Research (BMBF). Permission to sample arthropods with pitfall traps was granted by the Bezirksregierung Braunschweig.

References

ALDERWEIRELDT, M., J. HUBLE & M. POLLET, 1989: The Aerofauna of different woodland habitats of the »Lippensgoed-Bulskampveld« area (Beernem, Western Flanders, Belgium). – Biol. Jb. Dodonaea 57: 87–102.

BÖTTCHER, H., I. BAUER & H. EICHNER, 1981: Die Buchen-Waldgesellschaften des Fagion Sylvaticae im südlichen Niedersachsen. In: H. DIERSCHKE (ed.) Berichte der Internationalen Symposien der Internationalen Vereinigung für Vegetationskunde. Syntaxonomie (Rinteln 1980). – Cramer, Vaduz: 545–567.

DAMMANN, W., 1969: Physiologische Klimakarte Niedersachsens. – Neues Archiv f. Niedersachsen 18: 287–298.

DIERSCHKE, H., 1989: Kleinräumige Vegetationsstruktur und phänologischer Rhythmus eines Kalkbuchenwaldes. – Verh. Ges. Ökol. 17: 131–143.

DORNIEDEN, K., A. SÜHRIG, C. DÖRING & M. JUDAS, 1996: Analyse regionaler Verbreitungsmuster von Laufkäfern und Spinnen. – Artenschutzreport (Jena) 6: 46–49.

DÖRING, C., 1996: Zur Eignung GIS gestützer zoologischer Vorhersagemodelle für den Arten- und Biotopschutz. – Artenschutzreport (Jena) 6: 51–56.

HÄNGGI, A., E. STÖCKLI & W. NENTWIG, 1995: Lebensräume mitteleuropäischer Spinnen. – Misc. Faun. Helv. 4: 1–460.

JOCQUE, R., 1973: The spider fauna of adjacent woodland areas with different humus types. – Biol. Jb. Dodonaea 41: 153–179.

JUDAS, M., K. DORNIEDEN & C. DÖRING, 1998: Analysis of the regional distribution of epigeic arthropods 1. Distribution models for ground beetles. – Verh. Ges. Ökol. 28: #–#.

SCHAEFER, M., 1989: Die Bodentiere eines Kalkbuchenwaldes: Ein Ökosystemforschungsprojekt. – Verh. Ges. Ökol. 17: 203–212.

SÜHRIG, A., 1997: Die Spinnenfauna des Göttinger Waldes (Arachnida: Araneida). – Göttinger naturk. Schr. 4: 117–135.

THÖLE, R. & B. MEYER, 1979: Bodengenetische und ökologische Analyse eines Repräsentativareals der Göttinger Muschelkalkscholle als landschaftsökologische Planungsgrundlage. – Göttinger Bodenkundl. Ber. 59: 1–235.

Address

Axel Rothländer (arothla@gwdg.de)
Dipl.-Biol. Alexander Sührig
Dr. Michael Judas (mjudas@gwdg.de)
Prof. Dr. Matthias Schaefer (mschaef@gwdg.de)
Institut für Zoologie und Anthropologie
Abt. Ökologie
Berliner Str. 28
D-37073 Göttingen
Germany