Ecological Modelling 431 (2020) 109160

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Integration of landscape metric surfaces derived from vector data improves species distribution models

Olivia Ortner*, Gudrun Wallentin

Interfaculty Department of Geoinformatics – Z_GIS, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria

ARTICLE INFO

Keywords:
Maxent
Landscape metrics
Coronella austriaca
Zonal metrics
Tessellations

ABSTRACT

A species' distribution across the landscape is not random, but it is affected by distribution, size, abundance and connectivity of landscape patches. This spatial configuration of the landscape shapes ecological processes, for example the location of home ranges, migration routes and migration ability. Landscape metrics describe the configuration of a landscape quantitatively. While traditional approaches in habitat modelling only consider environmental attributes at a specific location, the integration of landscape metrics adds more functional information. In this paper we evaluated a method of directly incorporating a set of landscape metrics as covariates into a Maxent habitat model. Specifically, we used hexagons as statistical units for the calculation of landscape metrics. With this method also landscape metrics calculated with vector data sets can be used for SDM. We tested this approach for the smooth snake (*Coronella austriaca*) in the Austrian Alps. The experimental designs resulted in an improvement of the habitat models. Moreover, the results demonstrated the benefits of landscape metrics for the model outcomes at different scales.

1. Introduction

For effective conservation planning, it is fundamental to understand the factors that drive the distribution of species (Rosenzweig, 1995). Species distribution modelling is thus used in numerous studies that estimate the present and future habitat ranges of species for biodiversity research and conservation biology. The most commonly used predictors in species distribution models are environmental variables, like climate, vegetation, soil and elevation (Pulliam, 2000). While these variables describe the ecological niche at a certain location, they do not account for the spatial configuration and structure of the landscape. The spatial context, in which the species is embedded, has an important influence on ecosystem functions and thus also on habitat suitability and biodiversity (Walz, 2011). In other words, a species´ distribution is not only determined by its environmental niche, but also by functional relations in geographic space. We thus argue that species distribution models become more informative with using landscape metrics as covariates.

In order to adequately model the local distribution of a specific animal species, predictive variables need to be selected from an organism-centred perspective (Cushman et al., 2008; Li and Wu, 2004). For the analysis of landscape configuration, this implies the choice of an appropriate scale that matches dispersal processes of the species under study (Walz, 2011). Scale in this respect is characterised by the spatial resolution and extent, but also the thematic granularity of land-cover classes (Lam and Quattrochi, 1992; Turner et al., 2001). Species with higher space demand and higher mobility are influenced by a larger extent of landscape than small species with low mobility. So, spatial grain of habitat perception is a function of body size, what accords to the decision hierarchy concept of Holling (1992). To represent the actual driving forces for the distribution of a target species, resolution and extent thus need to reflect the size of the home ranges of this species (Holzkämper et al., 2006; Guisan and Thuiller, 2005). Moreover, the spatial and the thematic resolution of the model is limited by the quality of the underlying data to avoid pseudo-accuracy.

Landscape metrics are calculated with a set of algorithms that help to describe the spatial configuration of landscapes quantitatively. Therefore, these metrics can be an important - although mostly neglected - factor for species distribution modelling. Adding to the above conceptual considerations of scale, the use of landscape metrics as predictors in species distribution models also poses a methodological challenge if vector data is used. Schindler et al. (2013) and Turner et al. (2001) have demonstrated that the spatial scale and extent of a study area affects the performance of landscape metrics. Especially the response to changing extent is not consistent (Saura and Martinez-Millan, 2001). Small extents can cause the number of patches of the same class to drop below statistically meaningful sample sizes and thus lead to an unpredictable behaviour of metrics (Schindler et al., 2013).

* Corresponding author at: Seeparksiedlung 10/3, 5071 Wals, Austria.
E-mail addresses: ortner.olivia@gmx.at (O. Ortner), gudrun.wallentin@sbg.ac.at (G. Wallentin).

https://doi.org/10.1016/j.ecolmodel.2020.109160
Received 11 January 2020; Received in revised form 24 May 2020; Accepted 30 May 2020
0304-3800/

Some studies have used landscape metrics to complement species distribution models with additional information after the modelling process (e.g. Amici et al., 2015; Hasui et al., 2017; Hopkins 2009; Foltete et al., 2012), to predict species richness (e.g. Schindler et al., 2013) or to correlate landscape metrics with species occurrence (e.g. Ippoliti et al., 2013; Westphal et al., 2003). Another possibility to incorporate landscape metrics into the modelling process is the “moving window” approach, which is restricted to raster data sets (e.g. Hagen-Zanker, 2016; Santos et al., 2019).

In this research, we suggest a novel method to incorporate landscape metrics as continuous surfaces derived from vector data sets into a species distribution model to account for the spatial configuration of landscapes. So, this method enables the use of vector data sets for landscape metrics calculations and model building. Specifically, we (1) calculated landscape metrics for regular, hexagonal cells at different ecologically meaningful scales for the smooth snake (*Coronella austriaca*) in the Austrian Alps, (2) incorporated the landscape metrics as predictors into a species distribution model based on the Maxent algorithm to characterise the niche of *Coronella austriaca*, and (3) tested the approach against a model without landscape metrics.

2. Material and methods

The approach of integrating landscape metrics as continuous surfaces into the species distribution model of the smooth snake followed a 3-step workflow (Fig. 1): First, occurrence points were pre-processed to eliminate low precision and clustered points. Second, the most contributing landscape metrics types and scale levels were identified in an exploratory modelling phase. This step was done on a subset of the study area (“test areas”) to avoid computational overload. Third, species distribution models were computed with the selected environmental and landscape predictors for the whole study area.

2.1. Study species and area

The target species of this study was the smooth snake (*Coronella austriaca*). We chose the smooth snake, because the lead author is herpetologist with a profound knowledge of this colubrid. We thus had access to an abundant set of recent survey data and also could well interpret results. Although this snake is distributed across entire Europe, western Siberia and the Middle East (Völkl and Käsewieter, 2003) it is included in the European Council Directive 92/43/EEC of 21st of May 1992 Annex IV and has been evaluated as being in an “unfavourable state” in Central and Northern European countries (Čeirāns and Nikolajeva, 2017). *Coronella austriaca* is a rather small (mean length between 60–70 cm), non-venomous and secretive snake that is mainly threatened through habitat loss and fragmentation, which leads to extinction of populations and reduces the gene flow between persisting populations. This can cause degeneration of the remaining populations (Pernetta et al., 2011; Reading, 2012).

C. austriaca is one of the typical elements of the European cultural landscape and is very ductile in its habitat selection. It inhabits a wide spectrum of open and half-open landscapes and can be seen as xerothermophile species that sometimes also inhabits wet to alternating wet areas (Völkl and Käsewieter, 2003). All of these habitats are highly structured landscapes with adequate microhabitats like immature soil, dry grass, stone and rock and deadwood (Käsewieter, 2002).

The study area is located in the southern ranges of the Eastern Alps in Carinthia, a province of Austria. It is composed of 15.6% subalpine and alpine vegetation, 57.6% of different types of forest, 0.3% of wetlands, 19.4% agriculture and 7.1% of miscellaneous areas (Hartl et al., 2001) with an elevation between 384 m in the East and 3798 m in the West. For this province, a rich set of *Coronella austriaca* observation points was available together with a fine-scaled vegetation map at the scale of 1:50,000, which provided enough accuracy to detect relevant landscape patterns for this snake. Further, a representative test area was delimited, to explore the performance of landscape metrics surfaces with a computationally less demanding subset. The size and position of the test area was chosen because of the number of occurrences in the area and the fact, that it covers lower altitude regions as well as some subalpine areas (Fig. 2).

2.2. Occurrence data and environmental data

Occurrence data for *Coronella austriaca* was obtained from the Herpetofaunistic Database of the Museum of Natural History, Vienna[1] and from the Consortium for Nature Conservation, Klagenfurt.[2] Together, these databases held 1208 occurrence records. In the pre-processing, we extracted records that were collected between 1996 and 2016 to ensure a temporal match with the environmental layers. Further, metadata was used to filter out locations with a positional inaccuracy of 100 m or more. Finally, 129 occurrence records were left, 46 of which were located in the test area. In a second pre-processing step, we filtered the remaining occurrences spatially to reduce bias through spatial autocorrelation (Boria et al., 2014; Anderson and Gonzalez, 2011). This is necessary, because biased occurrence records can lead to overfitted model outputs in Maxent (Peterson et al., 2007), which means that the model is more complex than the real relationships between the included environmental variables and the species' niche (Peterson 2011). Therefore, only one randomly selected point was kept in occurrence clusters of 500 m distance. After spatial filtering, 94 occurrence points were left in the study area, of which 38 occurrences were located in the test area.

The current vegetation of Carinthia[3] with a detection scale of 1:50.000 served as the environmental layer for the computation of landscape metrics. This vegetation layer was further enriched with specific habitat information for *Coronella austriaca*, especially small water bodies (source: water body network of Carinthia[3]), wetlands (source: Map of the current vegetation of Carinthia (Hartl et al., 2001)), and alpine land-cover (source: the generalized land use of Carinthia[3]). The resulting layer contained 51 vegetation classes.

To calculate the different kinds of landscape metrics three hierarchical levels were used, two of them aggregations of the vegetation classes: (1) vegetation classes, (2) vegetation types, (3) land cover classes. The vegetation was aggregated into functional habitat types for *Coronella austriaca* like open and half-open landscape types and also wet to altering wet areas (Völkl and Käsewieter, 2003). Thus, three aggregation levels were available for further analysis: First, 51 vegetation classes at the level of plant associations e.g. “secondary spruce forest on carbonate ground”; second, 24 vegetation types, e.g. “spruce and mixed spruce forest”; and third, 7 land-cover classes. e.g. “forest”.

Finally, seven climatic layers with a resolution of 250 × 250 m were acquired for model-building: mean annual global radiation, average accumulated precipitation, average accumulated summer precipitation, mean snow cover duration, average start of snow cover, average end of snow cover, average equivalent temperature in July.[4]

2.3. Species distribution model

The algorithm that we used to model the distribution of the smooth snake was Maxent as implemented in Maxent GUI 3.4.1.[5] Maxent is a machine learning method, which today is one of the most frequently used algorithms to model species distribution (Phillips, 2004). It is a presence-background modelling method that associates known occurrences of a species with environmental data in the region of interest.

[1] Herpetofaunistische Datenbank des Naturhistorischen Museums, Wien.
[2] Arge NATURSCHUTZ, Klagenfurt
[3] https://data.gv.at/katalog/dataset
[4] https://data.gv.at/katalog/dataset
[5] http://biodiversityinformatics.amnh.org/open_source/maxent/

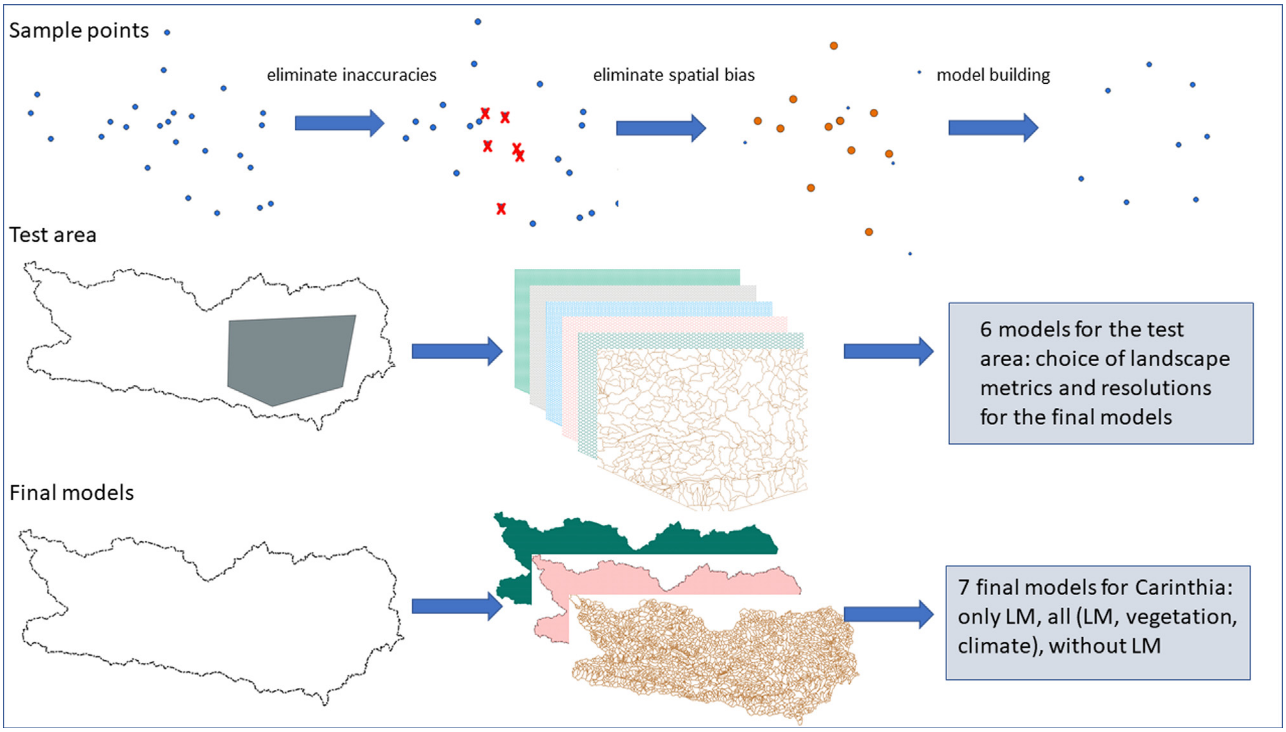

Fig. 1. Overview of the most important steps in building the model. (Sample points) First, the sample points (occurrence data of *Coronella austriaca*) were filtered, (Test area) second the most adequate landscape metrics and scale levels were identified with exploratory models in the test area and (Final models) third, the final species distribution models were built for Carinthia for three target scales.

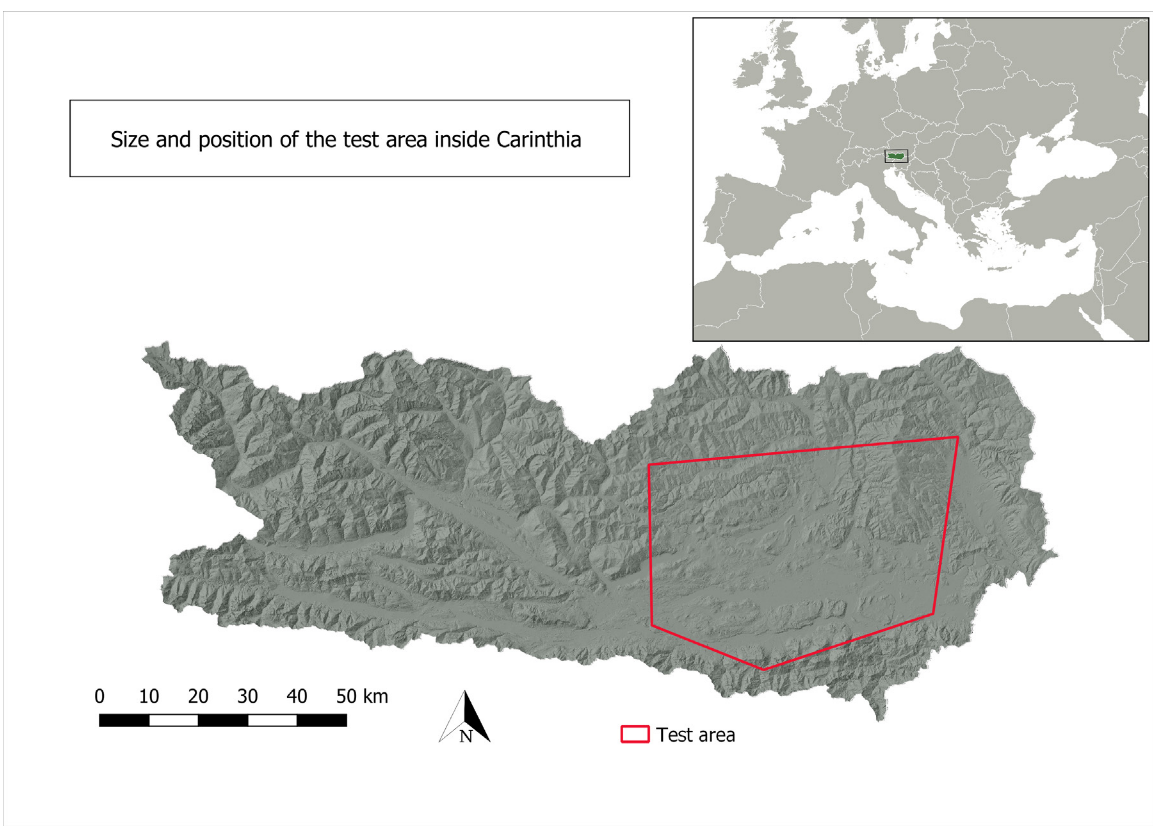

Fig. 2. The study area in the province of Carinthia, Austria. The area outlined in red is the test area, in which the test models were calculated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The resulting model extracts the ecological niche that the target species can inhabit in the study area and maps it onto geographic space. The algorithm consistently shows good results, especially for presence-only data (Merow et al., 2013; Elith et al., 2006).

In Maxent, there are different approaches how to select covariates in ecological modelling. One recommends reducing correlation between them to a minimum before starting the modelling process through correlation analysis, clustering analyses or another reduction method, because the complex features used by Maxent often produce highly correlated outputs. Reducing the covariates prior to model building should result in models that are better interpretable. This corresponds to the approach of treating Maxent as traditional statistical model (Renner and Warton, 2013). An alternative approach considers Maxent as machine learning method and lets the algorithm decide, which covariates to use for model building through regularization (Phillips et al., 2006). We followed the latter approach and did not filter the covariates before model building.

2.4. Regular tesselations

Landscape-level metrics are usually computed with respect to an entire landscape. The result for landscape metrics calculations is one number that characterises the whole study area. However, in order to incorporate the spatial configuration of a landscape as predictor into the Maxent modelling process, we needed to disaggregate the metrics into tessellation areas of landscape units. The size of landscape units thereby needs to account for the multiple spatial scales at which dispersal processes are operating. Landscape metrics were calculated for this small units and afterwards the tessellation areas were converted to continuous raster files.

To explore the computation of landscape metrics for different spatial scales, we tesselated the test area into multiple hexagonal grids of different cell sizes (5 ha, 10 ha, 15 ha, 25 ha, 35 ha per hexagon). The size of the statistical zones defines the scale of analysis. That should be considered because every analysed phenomenon can have a particular scale domain where it reveals (Levin, 1992; Turner et al., 1989). To build organism-centred models for the target species, the size of the hexagons was on the one hand adjusted to the habitat size requirements of *Coronella austriaca* populations, which ranges between 1 ha and >200 ha (Völkl and Käsewieter, 2003). On the other hand, we experimented with the size of the hexagons to get an insight on how the landscape metrics reacted. Boundary patches that extended over two hexagon zones or more were clipped. In addition to the hexagonal grids, we delimited the catchment areas of the test area with a mean zone size of 303 ha to serve as alternative, natural ecological units for the landscape metrics computation.

Finally, the landscape metrics tessellation layers were rasterised to a resolution of 100 m x 100 m. This was the finest resolution to avoid pseudo-accuracy in the modelling process, taking into account the resolution of the vegetation map and location inaccuracies for the observed snake occurrences.

2.5. Choice of candidate landscape metrics

Landscape metrics can be computed at three different levels, for individual patches (e.g. patch shape), patch classes with respect to a landscape (e.g. mean patch area or nearest neighbour distance), and the structure of the mosaic of patches in a landscape (e.g. fragmentation or connectivity). In this research, we were interested in the characterisation with respect to the entire landscape. Thus, we selected a set of class- and landscape-level metrics that were meaningful to describe the habitat of the smooth snake. In this exploratory phase, a large number of metrics was computed. As each model was spatially disaggregate, the calculations proved to be computationally demanding and we thus limited them to the representative test area described above.

For all statistical layers and the catchment areas, six types of landscape metrics were computed with the ZonalMetrics toolbox (Adamczky and Tiede, 2017) for ArcGIS (Esri, 2011) to quantify important habitat elements of *C. austriaca* (Table 1):

1 Area Metrics were calculated for open areas important for C. austriaca of the vegetation types. The result was the percent of the area of the whole statistical zone taken by the patch.
2 The Largest Patch Index was calculated for the plant associations. Result: percentage of the total area of the statistical zone taken by the largest patch.
3 Connectance Metrics for the habitable land cover, the maximum connectance distance was 500 m. The examined classes (covered with buildings, planted, grassland, and intensive grassland) were merged. The resulting values were the number of distinct connected classes, the percentage of patch area that lies within the range of connection to the statistical zone and the percentage of the connection zone between the patches in comparison to the statistical zone.
4 Contrast Metrics were calculated for the habitable land cover. The analysed classes (one at a time) were covered with buildings, intensive grassland, planted and grassland. The contrast classes were compact settlement, waterbodies and forest. The resulting value was the contrast index which is calculated as the percentage of the edge length of the focus classes shared with the contrast classes.
5 The Shannon Diversity Index was calculated for the plant associations. This Index considers the number of different patch types and their abundance.
6 The edge density was computed based on boundaries between vegetation classes and the transportation network. We dissolved the vegetation layer polygons to lines and merged the resulting layer with the transport network of Carinthia to generate a line kernel density surface. The decision for a bandwidth is a key step in kernel density estimation, depending on the smoothing of the resulting surface (Cai et al., 2013). As a rule of thumb ArcGIS (Esri, 2011) works with the rule of Silverman (Silverman, 1986), which is based on a quadratic kernel function. The first surface was calculated with the suggested bandwidth of 1741.11 m. However, to represent edge density for C. austrica in an appropriate way, the smoothing should not be too strong, because the effect of these edges does not expand to far away from the linear structures. Thus, three more surfaces with bandwidths of 500 m, 1000 m and 1500 m were created. Based on a comparison with the line data set we chose the surface with the 1000 m bandwidth and a resolution of 100 m x 100 m for the modelling process.

For landscape-level metrics like the Largest Patch Index or the Connectivity metrics, one single landscape metric surface resulted for each metric. Class-level metrics like the share of open areas resulted in multiple surfaces, one for each class. Thus, finally 26 landscape metrics surfaces at six resolutions resulted and were ready to be used as candidate predictors for further analyses in the test area.

2.6. Statistical selection of landscape metric surfaces

In this final pre-processing step, we singled out the landscape metric surfaces, which contributed most to the model, and determined the most adequate resolutions. To test respective models quantitatively against the validation data, we did a jack-knife evaluation with a random test percentage was 25% and a number of background points of 10,000, averaged over 20 iterations.

The assessment of model performance was based on two criteria: AUC (area under the ROC curve) and omission rate (OR). AUC reflects the discriminatory ability of a model. It is a measure for the ability, that a model ranks a random presence locality better than a random background point (Phillips et al., 2006), where higher values implicate better models. Omission rate (OR) quantifies overfitting. It ranges

Table 1
Overview over the 26 landscape metric surfaces that resulted from the computation of six different types of landscape metrics.

Landscape Metrics (LM)	Description	Level of thematic aggregation	Number of LM surfaces
Share of open areas	percent area occupied by a class	Vegetation types (only open landscapes)	8
Largest Patch Index	percent area occupied by the largest patch	Plant associations	1
Connectivity			
- Connectivity – classes	Number of connected habitable classes	Land-cover types	1
- Connectivity – habitat	Percent area occupied by a set of connected habitable patches	Land-cover types	1
- Connectivity – zone	Percent area of the connection zone between the patches	Land-cover types	1
Contrast	Percent edge length of the focus classes shared with the contrast classes.	Land-cover types (habitable vs. non-habitable)	12
Shannon Diversity Index	Considers the number of different patch types and their abundance	Plant associations	1
Edge Density	Line kernel density (Cai et al., 2013)	Plant associations (edges) and transport network	1

between 0 and 1, where lower numbers indicate better model performance (Shcheglovitova and Anderson, 2013). Although, the AUC based on presence-background data is an arguable absolute measure for the performance of models (Lobo et al., 2008; Warren and Seifert, 2011), it is adequate to compare models of single species in an identical study area (Peterson, 2011).

First, we determined the optimal resolution of the hexgrid in the test area. We computed models with all landscape metric surfaces for each of the six resolutions, i.e. the five hexgrid tesselations between 5 and 35 ha and the catchment units. The respective AUC and OR values were compared to identify the resolutions at which models performed best.

Second, we selected the landscape metric surfaces that were most important for the model predictions. As Maxent is a machine learning algorithm, we could make use of the regularisation procedure not only to avoid overfitting, but also to select the most contributing covariates (Phillips et al., 2006; Elith et al., 2011). Specifically, we used permutation importance to determine the relative importance of predictors. Permutation importance randomly alters the values of the covariate of interest among presence and background points. Subsequently, the drop in the AUC is measured and normalised to the percentage scale. The larger this drop, the more important is the covariate for the model prediction (Phillips et al., 2006). Only landscape metric surfaces with a permutation importance of 4% and more were kept as predictors for the final model. This threshold simplified the models to a reasonable amount of detail.

With two adjustments, we tried to avoid overfitting of the model: First, the model was constrained to linear and quadratic features. More complex feature combinations allow a better fine tuning, but are potentially very sensitive to a species' environmental tolerance, which in turn may lead to overfitted models (Shcheglovitova and Anderson, 2013). Second, we adapted the regularisation multiplier that controls the intensity of regularization across all features to a value of two. With this value the chance that the model is overfitted to bias or noise in the sample points is relatively low (Radosavljevic and Anderson, 2014).

Finally, we assessed how much model predictions differed in geographic space, even if their overall predictive performance was similar. The algorithms that we used to quantify geographic similarity were Schoeners *D* (Schoener, 1968) and *I* statistics (Warren et al., 2008). These performance indicators represent the difference between the normalised suitability scores per grid cell. The values range from 0 (no match) to 1 (identical models). While *I* often overestimates model similarity, Schoeners *D* is a more conservative measure (Rödder and Engler, 2011).

2.7. Final species distribution models

The final models were computed with climatic layers, the vegetation layer and the statistically most significant landscape metric surfaces. All predictors had the same raster resolution of 100 m x 100 m.

To assess the added value of incorporating landscape metrics as predictors, we computed 1) one model with all parameters, i.e. landscape metrics, climatic variables, and the vegetation layer, 2) one model with landscape metrics only, and 3) one model without landscape metrics. The latter null model served as a reference. The resulting species distribution maps (Fig. 3) were visualised as a binary prediction with 10 percentile training presence logistic threshold (Phillips and Dudík, 2008).

3. Results

3.1. Target resolutions

The model outputs for the different test area resolutions performed similarly well with respect to the resulting AUC and OR values. The values for the AUC ranged between 0.843 and 0.890 and for the OR between 0.346 and 0.369 for the 6 test area models (see Table 2 for the chosen model resolutions). Due to the similar performances across resolutions, we thus decided to select the three ecologically most meaningful scales for *Coronella austriaca* (Völkl and Käsewieter, 2003). One represents the population scale (5 ha), one the metapopulation scale (25 ha) and the catchment areas serve as natural ecological units.

3.2. Selection of landscape metrics

The individual contribution of the 26 test area landscape metrics on model predictions got marginally stronger, the coarser the hexgrid resolution was. The permutation importance for the 5 ha model contained five covariates with a zero percent permutation importance (contrast grassland-compact settlement, percentage of pioneer vegetation, contrast planted-compact settlement, percentage of subalpine vegetation, percentage of dwarf pine knee timber). The 25 ha model had four covariates with zero per cent permutation importance (contrast grassland-compact settlement, percentage of pioneer vegetation, percentage of dwarf pine knee timber, percentage of subalpine vegetation). The model with the highest zone size (catchment areas) only had one layer that did not contribute (percentage of subalpine vegetation). From the initial 26 landscape metrics surfaces only 6 of the 5 ha model, 5 of the 25 ha model and 7 of the catchment areas model had a 4% or more contribution to model building and thus were further used in the final models. Table 3 provides an overview of all variables that were used for the final models.

Despite some similarities in AUC and OR values between the different test area model results, we observed differences in geographic space. *Schoeners D* (Schoener, 1968) showed greater differences and less similarity than the *I* index (Warren et al., 2008). The highest *Schoeners D* value was 0.817 between the 5 ha and the 10 ha model. The lowest *Schoeners D* value, and therefore the highest difference showed the catchment area and the 5 ha model with 0.676. The average *D* value was 0.749. The highest *I* value was 0.971 between the 10 ha and the 15 ha model, the lowest *I* value was 0.907 between the 5 ha and the catchment area model. The average *I* value was 0.942.

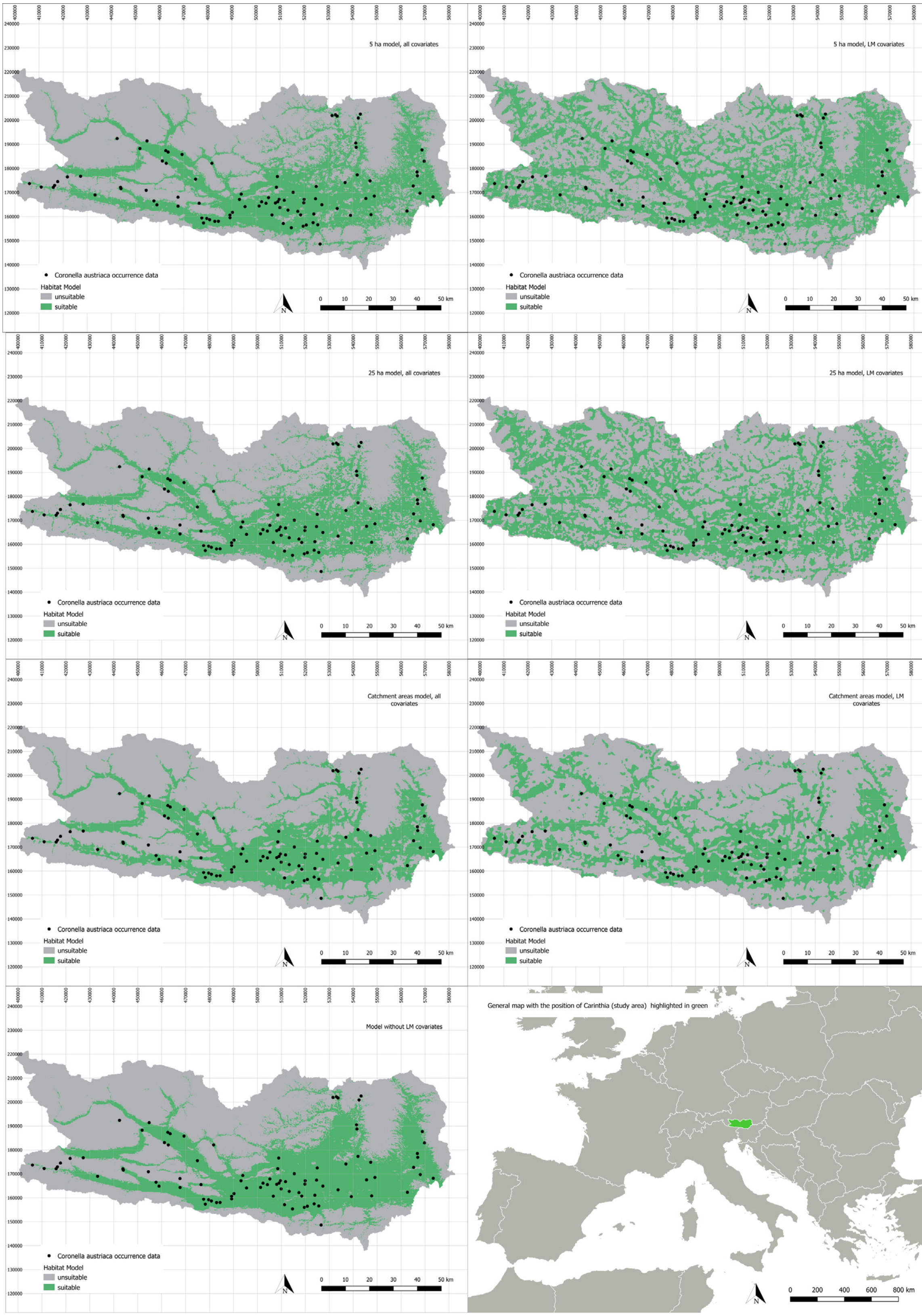

Fig. 3. presents the final species distribution maps for the province of Carinthia, based on binary predictions with a 10-percentile training presence logistic threshold plus an overview map.

3.3. The final models

The final models (Fig. 3) were calculated with the same settings as the test area models. For all three model resolutions the best values were shown by the models with all covariates: climatic, vegetation and landscape metrics surfaces (Table 4). The best value was 0.888 from the 25 ha model. The 5 ha model showed an AUC value of 0.882 and the catchment areas model AUC was 0.876. The AUC value of the model

Table 2
AUC and OR values for the three chosen resolutions of the test area models.

Resolution	AUC	OR
5 ha	0.8642	0.3686
25 ha	0.8899	0.3461
catchment areas	0.8854	0.3511

without landscape metrics surfaces was 0.857 and the values for the models with only landscape metrics were between 0.832 and 0.847, increasing with the size of the statistical zones. Considering the models without landscape metrics the model with the largest statistical surface units showed the best results. We additionally tried to fit the models by changing the settings (allow more feature classes and lower the regularisation multiplier), which increased the AUC values (between 0.859 for the 25 ha model with only landscape metrics and 0.928 for the 25 ha model with all covariates).

Compared to the test areas model runs, Schoeners *D* and *I* statistics both showed increased differences for all models in geographic space. The lowest *D* value with 0.610 appeared between the 5 ha model with landscape metrics only and the catchment areas model with all covariates. The highest *D* value was 0.850 between the 25 ha model with all covariates and the 5 ha model with all covariates. The average *D* value was 0.696. The *I* statistics again showed greater similarities between the models in geographic space. The lowest value was 0.848 between the 25 ha landscape metrics model and the model without landscape metrics. The highest value was 0.980 between the 25 ha surface with all covariates and the 5 ha surface with all covariates. The average *I* value was 0.907.

Table 4
AUC values for the final models. Models with all covariates include climatic, vegetation, and landscape metrics layers.

Model	AUC	Range AUC	Standard Deviation AUC
without landscape metrics	0.857	0.810–0.873	0,031
5 ha, all covariates	0.882	0.859–0.907	0,03
5 ha, only landscape metrics	0.832	0.791–0.881	0,038
25 ha, all covariates	0.888	0.859–0.919	0,033
25 ha, only landscape metrics	0.834	0.790–0.873	0,039
catchment areas, all covariates	0.876	0.845–0.907	0,03
catchment areas, only landscape metrics	0.847	0.799–0.888	0,035

The correlations between the covariates showed high values between the climatic layers from −0.72 between the average equivalent temperature in July and the average accumulated precipitation up to 0.99 between the mean snow cover duration and average start/end of snow cover, excluding the mean annual global radiation, where correlations were negligible with the other climatic layers. The landscape metrics showed very low correlations with the climatic layers, only the edge density and the average equivalent temperature in July sowed values up to 0.47. The landscape metrics surfaces among each other only showed higher correlations between the edge density, the Shannon Diversity Index and the largest patch index that ranged between −0.67 for the Shannon Diversity index and the largest patch index and 0.4 between the edge density and the largest patch index. The catchment areas model also showed higher correlations between the percentage of light building density and the edge density of 0.45. The vegetation layer

Table 3
The selection of covariates that were used for the final models.

Covariates Test Area Models	Covariates final models 5 ha - population level	25 ha - metapopulation level	catchment areas - natural ecological units
percentage of cultivated grassland		x	x
percentage of light building density	x		x
percentage of wetlands			
percentage of pastures and moutainious hay madows			
percentage of acre-grassland			x
percentage of dwarf pine knee timber			
percentage of pioneer vegetation on boulder and rocks			
percentage of subalpine and alpine grassland pastures			
Shannon diversity index	x	x	x
number of distinct connected classes			
percentage of patch area that lies within the range of connection			
percentage of the connection zone between the patches in comparison to the statistical zone			
Largest patch index	x	x	
contrast coverd with buildings- compact settlement			
contrast covered with buildings- water bodies	x		
contrast covered with buildings- forest			
contrast intensive grassland - compact settlement			
contrast intensive grassland - water bodies			
contrast intensive grassland - forest			x
contrast planted - compact settlements			
contrast planted - water bodies	x	x	
contrast planted - forest			
contrast grassland - compact settlement			
contrast grassland - water bodies			x
contrast grassland - forest			
edge density	x	x	x
vegetation layer	x	x	x
mean annual global radiation (kWh/m2)	x	x	x
average accumulated percipitation (mm)	x	x	x
average accumulated summer percipitation (mm)	x	x	x
mean snow cover duration (days)	x	x	x
average start of snow cover (day of the year)	x	x	x
average end of snow cover (day of the year)	x	x	x
average equivalent temperature in July (°C)	x	x	x

showed very low correlations between −0.25 and 0.33 with every other predictor. Again, the larger the statistical surfaces get, the stronger the correlations become.

4. Discussion

In contrast to the common assumption in species distribution modelling that the environmental niche is the main descriptor of distribution ranges (Elith and Leathwick, 2009), also geographic factors are influential. In this research, we were able to illustrate that the integration of landscape metrics as predictors can improve predictions of a species distribution model. among all seven distribution models for the smooth snake, the best AUC values resulted from combining all available predictor covariates (landscape structure and environmental variables). In contrast, the model with only environmental predictors performed worse, and the models with only landscape metric surfaces performed worst. These results confirmed our argument that a species' distribution is not only determined by its environmental niche, but also by functional relations in geographic space. Integration of such mechanistic aspects is especially important in non-equilibrium situations of expanding invasive species or climate-driven range shifts (Elith and Leathwick, 2009).

The most important landscape metrics covariates for *Coronella austriaca* were the Shannon diversity index, patch sizes and contrast metrics together with the egde density. That reflects some of the known habitat requirements of *Coronella austrica* like the preference of edge structures (Völkl and Käsewieter, 2003). When focusing on the contrast metrics (percentage of edge length of a focus calss shared with contrast classes), this also reflects the importance of edge effects for this species. The Shannon diversity index shows that the habitat suitability gets higher the more diverse the environment becomes. A more diverse environment also means more niches for different species and therefore more pray for *Coronella austriaca*. When examining the patch sizes *C. austriaca* seems to prefer regions with light building desity. This could also be an artefact because of the sampling intesity of the occurrence data. More occurence data was sampled in regions where people live and there still is enough room for snake habitats. So, areas with light building density may be overrepresented in the sample points. To put these results in a nutshell, the model confirms the known importand landscape traits for the target species.

4.1. Spatial structure

To account for the spatial context, we included "contextual indices" that summarised the characteristics within a spatial neighbourhood, like for example in Ferrier et al. (2002). Unlike these authors we were not interested in the surrounding ecological niches, but in the spatial structure itself. While the results of our research unambiguously show that distribution models improve with the integration of landscape metrics, this effect may depend on the species of interest. For example, Hasui et al. (2017) showed that the explanatory value of landscape metrics greatly varied between taxonomic groups. For small terrestrial mammals, the realtionship between habitat suitability and the landscape structure was confirmed by Amici et al. (2015).

Additionally, not only the scale of sampling is important, but also the grain of the underlying landscape. In order to obtain landscape metrics that are functionally relevant, it is important to select an adequate thematic granularity of vegetation classes with respect to the specific species under consideration. The resolution of the underlying data needs to be adapted to a specific species, its particular space demands and environmental requirements. This point often is neglected, especially in studies were multiple species are considered (e.g. Holzkämper et al., 2006; Schindler et al., 2013; Hasiu et al. 2017).

Also, the choice of adequate landscape metrics greatly depends on the species. Schindler et al. (2013) investigated the predictive power of landscape metrics for six very different taxa in more detail. They concluded, that patch shape, proximity, texture, diversity and patch size often were significant, whereas similarity or contrast metrics did not result in significant models. This partly is in accordance with the findings of our research, where also diversity (Shannon Diversity Index) and patch size (percentage of the area of the whole statistical zone taken by the patch) performed well as predictors in the model. In our research also contrast metrics showed good results. Although the target species of these studies are very different, it can be concluded that landscape metrics often contribute significantly to species distribution models.

Landscape metric predictor surfaces add distinctly new information to the environmental variables. We demonstrated in our research that despite a similar predictive power of environmental predictors versus spatial predictors, the Schoeners D analysis on map similarity showed that resulting distribution maps are of distinctly different nature. This is also confirmed by almost zero correlation between climatic and landscape layers, whereas the correlation between climatic layers is high. Interestingly, the correlation between various landscape metrics layers were also low. So each landscape metric predictor surface provides fundamentally new information.

4.2. Spatial scale

Another important aspect to consider for the incorporation of landscape metrics is the choice of scale. The spatial structure governs ecological processes at multiple scales. Our results confirm a better model performance for all "spatial" models. The mid-scaled cell resolution of 25 ha for the reference grid performed best. Accordingly, Václavík et al. (2012) observed an improvement in the distribution model of an invasive forest pathogen with the incorporation of spatial autocorrelation. Like in our model spatial dependences where highest in neighbourhoods of 200–400 m (12 to 50 ha). Also Schindler et al. (2013) found that the best performing size for the landscape metric reference areas for M. species was around 20 ha. Finally, Hasui et al. (2017) analysed broader-scale neighbourhoods (12.5 km^2), where the effects were not so clear. In their study, consideration of landscape metrics only had significant effects in 22% of the investigated target species. In summary, our research confirms the findings in the literature that the spatial structure of a landschape in the size of a few dozen hectars significantly impacts the occurrence of species. This seems a plausible size for a landscape reference grid, as it contains a representative sample of patches and enables fine scale modelling. This scale represents processes at the metapopulation level.

4.3. Recommendations and outlook

This research presents an approach to integrate the spatial structure of landscapes as predictors in species distribution models. The results indicate, that this approach holds promise and should be investigated further. It not only can enhance the predictive power of the model, but can also be helpful in identifying the most important landscape traits for the target species in the study area. However, the process to single out relevant landscape metrics and adequate scale levels is laborious and data hungry. Each decision had to be reflected and verified on the basis of the available data. Further research to better understand the role of scale and the adequacy of specific landscape metrics types may help to reduce this effort.

Credit author statement

Gudrun Wallentin:
Supervision, Conceptualization, Writing - Review&Editing
Olivia Ortner:
Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original draft, Visualization

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adamczyk, J., Tiede, D., 2017. ZonalMetrics - a Python toolbox for zonal landscape structure analysis. Comput. Geosci. 99, 91–99. https://doi.org/10.1016/j.cageo.2016.11.005. S.

Amici, V., Eggers, B., Geri, F., Battisti, C., 2015. Habitat suitability and landscape structure: a maximum entropy approach in a mediterranean area. Landsc. Res. 40 (2), 208–225. https://doi.org/10.1080/01426397.2013.774329. S.

Anderson, R.P., Gonzalez, I., 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. Ecol. Modell. 222 (15), 2796–2811. https://doi.org/10.1016/j.ecolmodel.2011.04.011. S.

Boria, R.A., Olson, L.E., Goodman, S.M., Anderson, R.P., 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecol. Modell. 275, 73–77. https://doi.org/10.1016/j.ecolmodel.2013.12.012. S.

Cai, X., Wu, Z., Cheng, J., 2013. Using kernel density estimation to assess the spatial pattern of road density and its impact on landscape fragmentation. Int. J. Geograph. Inf. Sci. 27 (2), 222–230. https://doi.org/10.1080/13658816.2012.663918. S.

Čeirāns, A., Nikolajeva, La., 2017. Habitat ecology of the smooth snake Coronella austriaca and its reptilian prey in the degraded bog with implications for artificial refuge surveys. Zool. Ecol. 27 (1), 19–29. https://doi.org/10.1080/21658005.2016.1252125. S.

Cushman, S.A., McGarigal, K., Neel, M.C., 2008. Parsimony in landscape metrics: strength, universality, and consistency. Ecol. Indic. 8 (5), 691–703. https://doi.org/10.1016/j.ecolind.2007.12.002. S.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29 (2), 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x. S.

Elith, J., Leathwick, J.R, 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Syst. 40 (1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159. S.

Elith, J., Phillips, S.J, Hastie, T., Dudik, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib (Divers. Distrib.) 17, S43–S57.

ESRI, 2011. ArcGis Desktop: Release 10. Redlands. CA: Environmental Systems Research Institute.

Ferrier, S., Watson, G., Pearce, J., Drielsma, M., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. Biodivers. Conserv. 11 (12), 2275–2307. https://doi.org/10.1023/A:1021302930424. S.

Foltete, J.-C., Clauzel, C., Vuidel, G., Tourant, P., 2012. Integrating graph-based connectivity metrics into species distribution models. Landsc. Ecol. 27, 557–569. https://doi.org/10.1007/s10980-012-9709-4. S.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8 (9), 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x. S.

Hagen-Zanker, A., 2016. A computational framework for generalized moving windows and its application to landscape pattern analysis. Int. J. Appl. Earth Obs. Geoinf. 44, 205–216.

Hartl, H.; Stern, R.; Seger, M. (2001): Karte der aktuellen Vegetation von Kärnten (M 1: 150.000);. = map of the actual vegetation of Carinthia, Austria: das Vegetationsgefüge einer inneralpinen Region im Süden Österreichs. Klagenfurt: Naturwissenschaftlicher Verein für Kärnten.

Hasui, É., Silva, V.X, Cunha, R.G.T., Ramos, F.N., Ribeiro, M.C., Sacramento, M., et al., 2017. Additions of landscape metrics improve predictions of occurrence of species distribution models. J. For. Res. 28 (5), 963–974. https://doi.org/10.1007/s11676-017-0388-5. S.

Holling, C.S., 1992. Cross-scale morphology, geometry, and dynamics of ecosystems. Ecol. Monogr. 62 (4), 447–502. https://doi.org/10.2307/2937313. S.

Holzkämper, A., Lausch, A., Seppelt, R., 2006. Optimizing landscape configuration to enhance habitat suitability for species with contrasting habitat requirements. Ecol. Modell. 198 (3–4), 277–292. https://doi.org/10.1016/j.ecolmodel.2006.05.001. S.

Hopkins, R.L., 2009. Use of landscape pattern metrics and multiscale data in aquatic species distribution models: a case study of a freshwater mussel. Landsc. Ecol. 24 (7), 943–955. https://doi.org/10.1007/s10980-009-9373-5. S.

Ippoliti, C., Gilbert, M., Vanhuysse, S., Goffredo, M., Satta, G., Wolff, E., Conte, A., 2013. Can landscape metrics help determine the Culicoides imicola distribution in Italy? Geospat. Health 8 (1), 267–277. https://doi.org/10.4081/gh.2013.72. S.

Käsewieter, 2002. Ökologische Untersuchungen an der Schlingnatter (Coronella austriaca, Laurenti 1768). Dissertation an der Universität Bayreuth.

Lam, N.S.-N., Quattrochi, D.A, 1992. On the issues of scale, resolution, and fractal analysis in the mapping sciences. Profession. Geograph. 44 (1), 88–98. https://doi.org/10.1111/j.0033-0124.1992.00088.x. S.

Levin, S.A., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73 (6), 1943–1967. https://doi.org/10.2307/1941447. S.

Li, Harbin, Wu, Jianguo, 2004. Use and misuse of landscape indices. Landsc. Ecol. 19 (4), 389–399. https://doi.org/10.1023/B:LAND.0000030441.15628.d6. S.

Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecol. Biogeograp. 17 (2), 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x. S.

Merow, C., Smith, M.J, Silander, J.A, 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36 (10), 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x. S.

Pernetta, A.P., Allen, J.A., Beebee, T.J.C., Reading, C.J., 2011. Fine-scale population genetic structure and sex-biased dispersal in the smooth snake (Coronella austriaca) in southern England. Heredity 107 (3), 231–238. https://doi.org/10.1038/hdy.2011.7. S.

Peterson, A.T., 2011. Ecological Niches and Geographic Distributions. Princeton University Press, Princeton N.J. (Monographs in population biology, 49).

Peterson, A.T., Papeş, M., Eaton, M., 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30 (4), 550–560. https://doi.org/10.1111/j.0906-7590.2007.05102.x. S.

Phillips, S.J., 2004. A maximum entropy approach to species distribution modelling. In: *Proceedings of the Twenty-First International Conference on Machine Learning* 2004, pp. 655–662 S.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190 (3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026. S.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. https://doi.org/10.1111/j.0906-7590.2007.5203.x. 080328142746259-.

Pulliam, H.R., 2000. On the relationship between niche and distribution. Ecol. Lett. 3, 349–361 doi: https://doi.org/10.1046/j.1461-0248.2000.00143.x.

Radosavljevic, A., Anderson, R.P, 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. J. Biogeogr. 41 (4), 629–643. https://doi.org/10.1111/jbi.12227. S.

Reading, C.J., 2012. Ranging behaviour and home range size of smooth snake inhabiting lowland heath in southern England. Herpetol. J. 22, 241–247 S.

Renner, I.W., Warton, D.I., 2013. Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. Biometrics 69 (1), 274–281. https://doi.org/10.1111/j.1541-0420.2012.01824.x. S.

Rosenzweig, M.L., 1995. Species Diversity in Space and Time. Cambridge University Press DOI: https://doi.org/10.1017/CBO9780511623387.

Rödder, D., Engler, J.O., 2011. Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. Global Ecol Biogeograp. 20 (6), 915–927. https://doi.org/10.1111/j.1466-8238.2011.00659.x. S.

Santos, J.P., Sobral-Souza, T., Brown, K.S., Vancine, M.H., Ribeiro, M.C., Freitas, A.V.L., 2019. Effects of landscape modification on species richness patterns of fruit-feeding butterflies in Brazilian Atlantic Forest. Divers. Distrib. 196–208. https://doi.org/10.1111/ddi.13007.

Saura, S., Martinez-Millan, J., 2001. Sensitivity of landscape pattern metrics to map spatial extent. Photogramm. Eng. Remote. Sens. 67, 1027–1036 S.

Schindler, S., Wehrden, H.v., Poirazidis, K., Wrbka, T., Kati, V., 2013. Multiscale performance of landscape metrics as indicators of species richness of plants, insects and vertebrates. Ecol. Indic. 31, 41–48. https://doi.org/10.1016/j.ecolind.2012.04.012. S.

Schoener, T.W., 1968. The anolis lizards of Bimini. Resource partitioning in a Complex Fauna. Ecology 49 (4), 704–726. https://doi.org/10.2307/1935534. S.

Shcheglovitova, M., Anderson, R.P, 2013. Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. Ecol. Modell. 269, 9–17. https://doi.org/10.1016/j.ecolmodel.2013.08.011. S.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Springer, Boston, MA s.l.US (Monographs on Statistics and Applied Probability). Online http://dx.doi.org/10.1007/978-1-4899-3324-9.

Turner, M.G., Dale, V.H, Gardner, R.H., 1989. Predicting across scale: theory development and testing. Landsc. Ecol. 3, 245–252 S.Online http://dx.doi.org/10.1007/bf00131542.

Turner, M.G., Gardner, R.H., O'Neill, R.V., 2001. Landscape Ecology in Theory and practice. Pattern and Process. Springer, New York Online. http://lib.myilibrary.com?id=18739.

Václavík, T., Kupfer, J.A, Meentemeyer, R.K., 2012. Accounting for multi-scale spatial autocorreltation improves performance of invasive species distribution modelling (iSDM). J. Biogeogr. 39 (1), 42–55. https://doi.org/10.1111/j.1365-2699.2011.02589.x. S.

Völkl, W., Käsewieter, D., 2003. Die Schlingnatter. Ein heimlicher Jäger. Bielefeld: Laurenti-Verl. 6 Zeitschrift für Feldherpetologie Beiheft.

Walz, U., 2011. Landscape structure, landscape metrics and biodiversity. Living Rev. Landsc. Res. 2011 (5) Online. http://www.livingreviews.org/lrlr-2011-3.

Warren, D.L., Glor, R.E., Turelli, M., 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62 (11), 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x. S.

Warren, D.L., Seifert, S.N., 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol. Appl. 21 (2), 335–342 S.

Westphal, M.I., Field, S.A., Tyre, A.J., Paton, D., Possingham, H.P., 2003. Effects of landscape pattern on bird species distribution in the Mt. Lofty Ranges, South Australia. Landsc. Ecol. 18 (4), 413–426. https://doi.org/10.1023/A:1026115807529. S.