



RESEARCH ARTICLE

Optimization of spatial scale, but not functional shape, affects the performance of habitat suitability models: a case study of tigers (*Panthera tigris*) in Thailand

Eric Ash · David W. Macdonald · Samuel A. Cushman · Adisorn Noochdumrong ·
Tim Redford · Zanita Kaszta

Received: 20 March 2020 / Accepted: 29 August 2020
© The Author(s) 2020

Abstract

Context Species habitat suitability models rarely incorporate multiple spatial scales or functional shapes of a species' response to covariates. Optimizing models for these factors may produce more robust, reliable, and informative habitat suitability models, which can be beneficial for the conservation of rare and endangered species, such as tigers (*Panthera tigris*).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10980-020-01105-6>) contains supplementary material, which is available to authorized users.

E. Ash · D. W. Macdonald · S. A. Cushman · Z. Kaszta
Wildlife Conservation Research Unit, Department of
Zoology, The Recanati-Kaplan Centre, University of
Oxford, Tubney House, Tubney, Oxon OX13 5QL, UK

E. Ash · T. Redford
Freeland Foundation, 92/1 Soi Phahonyothin 5,
Phahonyothin Road, Phaya Thai, Bangkok 10400,
Thailand

S. A. Cushman (✉)
Rocky Mountain Research Station, United States Forest
Service, Flagstaff, AZ, USA
e-mail: scushman@fs.fed.us

A. Noochdumrong
Ministry of Natural Resources and Environment, 92 Soi
Phahonyothin 7, Phahonyothin Road, Phaya Thai, Phaya
Thai, Bangkok 10400, Thailand

Objectives We provide the first formal assessment of the relative impacts of scale-optimization and shape-optimization on model performance and habitat suitability predictions. We explored how optimization influences conclusions regarding habitat selection and mapped probability of occurrence.

Methods We collated environmental variables expected to affect tiger occurrence, calculating focal statistics and landscape metrics at spatial scales ranging from 250 m to 16 km. We then constructed a set of presence-absence generalized linear models including: (1) single-scale optimized models (SSO); (2) a multi-scale optimized model (MSO); (3) single-scale shape-optimized models (SSSO) and (4) a multi-scale- and shape-optimized model (MSSO). We compared performance and resulting prediction maps for top performing models.

Results The SSO (16 km), SSSO (16 km), MSO, and MSSO models performed equally well ($AUC > 0.9$). However, these differed substantially in prediction and mapped habitat suitability, leading to different ecological understanding and potentially divergent conservation recommendations. Habitat selection was highly scale-dependent and the strongest relationships with environmental variables were at the broadest scales analysed. Modelling approach had a substantial influence in variable importance among top models.

Conclusions Our results suggest that optimization of the scale of resource selection is crucial in modelling tiger habitat selection. However, in this analysis,

shape-optimization did not improve model performance.

Keywords Functional form · Habitat selection · *Panthera tigris* · Scale-optimization · Shape-optimization · Dong Phayayen–Khao Yai Forest Complex

Introduction

Much of modern ecological analysis is underpinned by ecological niche theory (Hutchinson 1957). At the center of ecological inquiries is the organism, with its distribution, abundance, and survival dictated by its niche, defined as a function of limited resources within an n -dimensional hypervolume of continuous space (Hegel et al. 2010; Blonder 2018). The proliferation of modern analytical tools (Elith et al. 2006; Blonder 2018) and increased computational power to quantify these fundamental and realized niches has advanced the field of landscape ecology and enabled more powerful habitat selection modelling (Hegel et al. 2010). However, such assessments, even for well-studied species, are rarely straightforward (Mayor et al. 2009).

Gradients of factors along niche hypervolume axes can exhibit complex patterns across a continuum of spatial scales (Cushman 2007; Hegel et al. 2010), with organism responses to its environment is simultaneously influenced by both fine- and broad-scale environmental attributes (Wiens 1976; Johnson et al. 1992; Levin 1992). No single spatial scale is typically sufficient to elucidate organism-habitat relationships. Thus, investigations into habitat selection should be conducted at multiple, ecologically-relevant spatial scales (Wiens 1989; Goodwin and Fahrig 1998; McGarigal et al. 2016). In addition, organism responses to environmental and other factors, may not necessarily be linear. Austin et al. (1990) provide evidence of complex, non-linear organism responses in habitat studies, a result reflected in a number of other studies on species-habitat relationships (Wasserman et al. 2012; Hebblewhite et al. 2014; Timm et al. 2016; Bosco et al. 2019; Macdonald et al. 2019). However, related assumptions pertaining to these relationships have largely been untested.

The importance of incorporating spatial scale into modelling of species–habitat relationships is well-established (McGarigal and Cushman 2002; McGarigal et al. 2016). Failing to do so can undermine the performance of habitat selection models and their interpretation (e.g., Wasserman et al. 2012; Mateo-Sanchez et al. 2014; Shirk et al. 2014), potentially leading to errors of inference and application (McGarigal and Cushman 2002). Optimized multiple-scale approaches (see definitions by McGarigal et al. (2016)) have been used to model habitat selection, limiting factors, and threats for a number of species (e.g. Thompson and McGarigal 2002; Toews 2011; Gonthier et al. 2014; Wan et al. 2017), and typically out-perform single-scale models in predictive or explanatory power (Kanagaraj et al. 2011; Wasserman et al. 2012; Mateo-Sanchez et al. 2014; Timm et al. 2016; McGarigal et al. 2016). It is notable, that despite the increased consideration of scale in habitat suitability studies since the publication of the influential McGarigal et al. (2016) review paper, few studies have formally applied scale-optimization to evaluate the effects of landscape components on species habitat suitability at multiple scales.

Conversely, the extent to which the functional shape of a species' response to covariates (i.e. functional form) affects model performance has largely been overlooked. Assuming a single response type for data being modelled may fail to account for different forms of relationship, an assumption that may not be appropriate even for modelling simple ecological associations (Oksanen and Minchin 2002). Indeed, the functional form of species–habitat relationships can vary across habitat characteristics (Redfern et al. 2006) and scales (Wan et al. 2017). Inclusion of variable-specific, non-linear functional relationships may strengthen the relationship between dependent and independent variables, improving representation of covariates in top models that may otherwise be obscured and, potentially, improving explanatory power (Carle 2006).

Generalized linear models (GLMs) have been the most widely-used approach in habitat modelling studies, given their intuitiveness and flexibility (McGarigal et al. 2016). GLMs assume a linear relationship between link function-transformed predictor variables and explanatory variables (Nelder and Wedderburn 1972). However, variables may have unique, non-linear response shapes. The vast majority

of habitat relationships studies have assumed linear relationships. In the relatively few cases in which different functional forms have been explicitly investigated, they have typically been restricted to quadratic relationships (Mateo-Sanchez et al. 2014; Mashintonio et al. 2014; Dzialak et al. 2015; Bosco et al. 2019; Macdonald et al. 2019). A few studies have examined other, more complex relationships (Bar-Massada et al. 2011; Fisher et al. 2011; Mateo-Sánchez et al. 2015; DeVoe et al. 2015; Shirk et al. 2018), although this appears not to have widespread application. Importantly, no published study, to our knowledge, has formally compared the performance and interpretation of models with and without functional shape-optimization.

The effect of scale- and shape-optimization in quantifying species niches and, by extension, species-habitat relationships merits investigation. Our goal in this paper is to explicitly evaluate the effects of scale- and shape-optimization on modelling habitat selection, using a globally important and understudied tiger population in Thailand as a case study. Tigers are an ideal species in which to investigate this, given large home ranges (Smith 1993; Sunquist 2010), potential influence of broad-scale environmental variation (Krishnamurthy et al. 2016; Reddy et al. 2017), and high conservation priority (Lynam and Nowell 2011; Goodrich et al. 2015). Using long-term camera-trap studies, we evaluate how shape- and scale-optimization affect predictive power and interpretation of models in comparison with non-optimized models. In doing so, we also evaluate how covariates influence habitat selection of tigers and the degree of scale- and shape-dependency.

We test four central hypotheses. First, we expect that habitat selection is highly scale-dependent, with clear patterns of relationship between the scale of analysis and strength of relationship between occurrence and environmental variables. Second, we expect that, for most variables, the relationship with tiger occurrence will be strongest at broad spatial scales, reflecting large home range sizes and high sensitivity to human influences. Third, we expect the multi scale-optimized model to outperform any of the single-scale optimized models, given evidence from previous studies. Lastly, we expect that shape-optimized models will outperform non-optimized models, given each variable is included with its independent shape of functional relationship with species occurrence.

Methods

Study site

The Dong Phrayayen–Khao Yai Forest Complex (DPKY) covers 6155 km² in eastern Thailand (Fig. 1). It consists of five protected areas—Khao Yai National Park (KYNP), Thap Lan National Park (TLNP), Pang Sida National Park (PSNP), Ta Phraya National Park (TPNP), and Dong Yai Wildlife Sanctuary (DYWS). The complex is relatively mountainous with elevations ranging from 100 m to 1351 m a.s.l. It contains a number of forest types, but primarily consists of mixed evergreen (77%) and mixed dipterocarp/deciduous primary and secondary forest (6%). The complex is biologically diverse, supporting 112 mammalian species, 392 birds, 200 reptiles/amphibians, and 2500 plant species (DNP 2004), and is on the UNESCO World Heritage List in recognition of its outstanding natural and conservation value (UNESCO 2017). Currently, DPKY is almost completely surrounded by villages and human infrastructure.

Camera-trap surveys and data processing

From 2008 to 2017, we conducted camera-trap surveys throughout DPKY to investigate tiger presence in each of its five protected areas (Ash et al. 2020). Survey coverage and intensity increased during this period as funding, camera-traps, and related resources increased, covering diverse habitats across DPKY. Within these areas, cameras were placed to maximize detection of tigers by prioritizing camera placement in areas with previous tiger or prey records and identifying topographic (e.g. ridges, river valleys) or other features (e.g. roads, trails) likely used by tigers (Karanth and Chundawat 2002; Sunarto et al. 2012; Barber-Meyer et al. 2013). To reduce spatial autocorrelation, if more than one camera was present within a 300 m radius, we selected the camera with the highest number of tiger detections or, if no tigers were detected, by greatest survey effort (camera-trap nights [CTN]). A more detailed description of surveys is provided in Ash et al. (2020) and supplementary online materials (Online Resource 1).

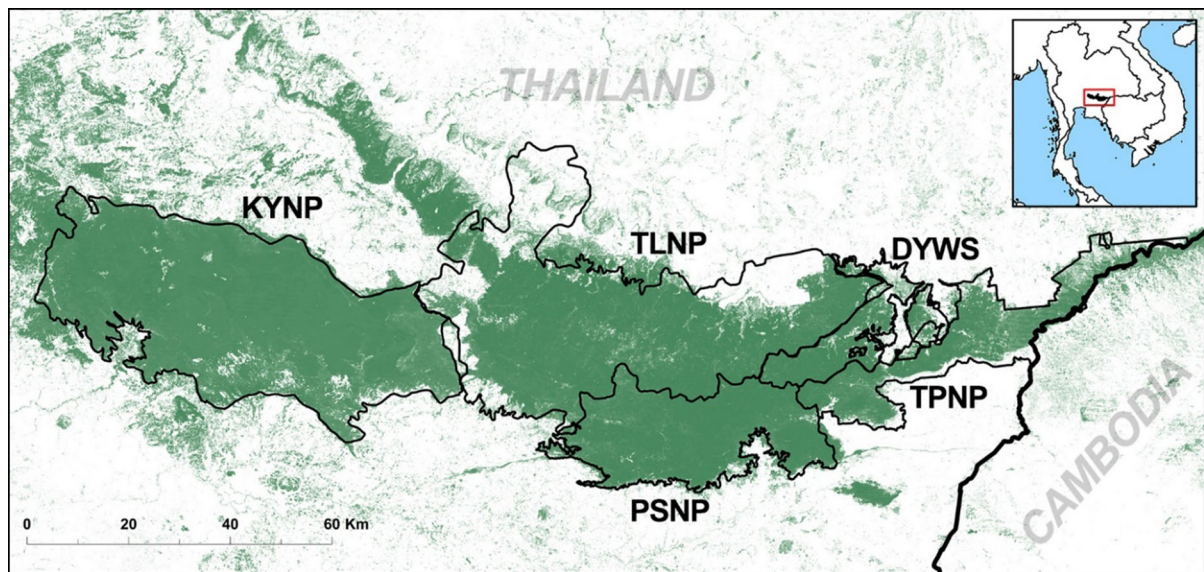


Fig. 1 Map of the Dong Phrayayen–Khao Yai forest complex (DPKY) and its five protected areas ($\sim 14^{\circ} 00'$ to $14^{\circ} 33'$ N and $\sim 101^{\circ} 05'$ to $103^{\circ} 14'$ E)—Dong Yai Wildlife Sanctuary (DYWS), Khao Yai National Park (KYNP), Pang Sida National Park (PSNP), Thap Lan National Park (TLNP), and Ta Phraya

National Park (TPNP). Map generated with ArcGIS 10.3.1 (Environmental Systems Research Incorporated, ESRI, Redlands, CA, USA, 2011) with forest cover derived from Hansen et al. (2013)

Predictor variables

We compiled a set of nine environmental predictor variables based on previous studies of tigers throughout their range (Wibisono et al. 2011; Sunarto et al. 2012; Ngoprasert et al. 2012; Kafley et al. 2016; Thapa and Kelly 2017). These primary variables were further transformed into more biologically informative predictor variables (Table 1) using class and landscape level spatial statistics (McGarigal et al. 2012). Variables were prepared in 30 m resolution to ensure representation of fine-scale variations in the environment.

We used SRTM 1 (Shuttle Radar Topography Mission) arc-second global elevation data (Jarvis et al. 2008), which we further transformed into several variables measuring topographic heterogeneity using the Geomorphometry & Gradient Metrics Toolbox (Evans et al. 2014) in ArcGIS 10.3.1 (Environmental Systems Research Incorporated, ESRI, Redlands, CA, USA, 2011). These included terrain roughness index (degree of elevation difference between adjacent cells), slope position (surface to area ratio) and compound topographic index (CTI). CTI is a measure of flow accumulation (Beven and Kirkby 1979), with low CTI values associated with ridges and

mountaintops and high values associated with large, low-lying drainage areas.

Percent forest cover (Hansen et al. 2013) was reclassified into three classes, non-forest (0–20%), open forest (20–40%) and closed forest (> 40%). Furthermore, the global land cover map (European Space Agency 2015) was classified into eight study-relevant area classes. Since some forest types (Royal Forestry Department 2000) were similar or under-represented in our study area, we reclassified eight forest types from an original 12 classes. Streams and rivers were included, but were reclassified to only include perennial rivers/streams (Royal Forestry Department 2000).

To test associations at different spatial scales, we transformed variables into seven multi-scale predictor variables (250 m, 500 m, 1 km, 2 km, 4 km, 8 km, and 16 km). For continuous variables, we generated focal statistics (focal mean and standard deviation) in ArcGIS using a moving circular window with radius equal to each scale. For categorical variables, we calculated landscape metrics in FRAGSTATS (McGarigal et al. 2012) using moving windows for each spatial scale. For class-level statistics, we calculated area-weighted mean radius of gyration (GYR, also known as correlation length), which is a measure

Table 1 Variables included in univariate model selection, metrics, and data sources

Variable	ID	Metric	Data Source
Elevation	DEM	Focal mean & SD	Jarvis et al. (2008; http://srtm.csi.cgiar.org)
Protected area boundary	BOUND	Focal mean	Royal Forestry Department (2000)
Forest class: 1-Non-forest (0–20%); 2-Open canopy (20–40%); 3-Closed canopy (> 40%)	TC	PLAND & GYR	Reclassified from Hansen et al. (2013)
Forest type: 1-Evergreen; 2-Mixed Deciduous; 3-Dry Dipterocarp; 4-Reforested Areas; 5-Bamboo Forest; 6-Secondary Forest/Old Clearing; 7-Grassland/Scrubland	FT	PLAND & GYR	Reclassified from Royal Forestry Department (2000)
Terrain Roughness Index (TRI)	TRI	Focal mean & SD	Derived from Jarvis et al. (2008; http://srtm.csi.cgiar.org) and Evans et al. (2014)
Slope position	SLOPE	Focal mean & SD	Derived from Jarvis et al. (2008; http://srtm.csi.cgiar.org) and Evans et al. (2014)
Streams/Rivers	STRC1	Focal mean	Reclassified from Royal Forestry Department (2000)
Compound topographic index (i.e., flow accumulation)	CTI	Focal mean & SD	Derived from Jarvis et al. (2008; http://srtm.csi.cgiar.org) and Evans et al. (2014)
Land cover classes: 1-Cropland; 2-Mosaic cropland; 3-Forest/Natural vegetation; 4-Shrubland/grassland; 5-Sparse vegetation & bare areas; 6-Urban Areas; 7-Water	LC	PLAND, GYR, CWED, PD, AI	Reclassified from European Space Agency (2015; www.esa-landcover-cci.org/?q=node/175)

Landscape statistics include focal mean and standard deviation (SD). Landscape metrics include percentage of landscape (PLAND), area weighted mean of radius of gyration/correlation length (GYR), contrast-weighted edge density (CWED), patch density (PD), and aggregation index (AI)

of mean potential travel distance within a habitat patch, or habitat extent, and percentage of habitat within the focal landscape (PLAND). At the landscape-level we calculated contrast-weighted edge density (CWED), which quantifies the influence of habitat edges via the degree of contrast between adjacent patches. We also calculated patch density (PD), a measure of the subdivision of habitat within the landscape (i.e. number of patches divided by area). Lastly, we calculated aggregation index (AI), which quantifies the degree of dispersion of habitat patches within the landscape. In total, we prepared 47 variables to investigate how spatial composition and configuration of the landscape at various scales influences tiger detections (Table 1).

Data analysis

To assess the influence of scale- and shape-optimization on model performance, we compared four approaches: (1) Single-scale optimized models

(SSO); (2) A multi-scale optimized model (MSO); (3) Single-scale shape-optimized models (SSSO); and (4) A multi-scale-and shape-optimized model (MSSO).

Model 1: Single-scale optimized model approach

To provide a baseline for assessing the effects of optimization on model performance, we developed a suite of seven independent models, one for each scale (250 m to 16 km), in which no optimization steps were applied (single-scale optimized [SSO]). This corresponds to what McGarigal et al. (2016) termed *pseudo-optimized single-scale* (ms4). For these models, the number of variables was reduced through a series of filtering steps. (1) Using the *glmer* function in the *lme4* package in R (Bates et al. 2014; R Development Core Team 2017), we produced a univariate generalized linear model (GLM) with a logit link function for each variable at each scale. (2) We required that $p < 0.05$ for each variable. (3) For all

variables with a Pearson's correlation coefficient $rl > 0.6$, we selected the variable with the lowest AICc value in each univariate GLM (Burnham and Anderson 2002).

These filtering steps resulted in the inclusion of 11 to 19 variables, including survey effort (CTN), among each of the seven scales (Online Resource 2, Table 2). To find the most parsimonious model for each spatial scale we used the 'biostats' package in R (McGarigal 2018), a set of functions which (1) produced models for all combinations of the standardized predictor variables, subsetting the best performing models ranked by $\Delta AICc$ (defined in our models as $\Delta AICc < 2$); (2) averaged a final set of models and calculated goodness-of-fit statistics, including proportion deviance explained for each candidate model; and (3) ranked variable importance in the final model with averaged coefficients based on Akaike's model weight (w_i).

Model 2: Multi-scale optimized model approach

For the multi-scale optimized (MSO) model, we selected the scale of best fit for each variable prior to the filtering steps and inclusion in the final model. For each variable, we developed a univariate GLM for each scale and selected the best performing (i.e. most optimal) scale by lowest AICc value. As in other models, we required $p < 0.05$ for each variable and

selected among correlated variables (Pearson's coefficient $rl > 0.6$) by lowest AICc value. Scale-specific variables were then included in a multivariate GLM using the 'biostats' set of functions (McGarigal 2018) to subset best performing models, calculate goodness-of-fit statistics, and rank variable importance in the final model. This approach is what McGarigal et al. (2016) termed (*pseudo-optimized multiple scales (ms5)*), as we optimized our model based on a range of scales determined a priori rather than selecting among an unconstrained range of continuous scales.

Model 3: Single-scale and shape-optimized model approach

To investigate the effects of shape-optimization on model performance, we developed seven independent single-scale shape-optimized (SSSO) models, one for each scale (250 m to 16 km), in which variables were included at an optimal functional form. To determine candidate functional forms to investigate, we plotted nonlinear splines using *plsmo* in the R package *Hmisc* (Harrell 2018) between variables and binomial tiger detections. Among resulting plots, we identified four potential functional forms (quadratic, logarithmic, exponential, and negative exponential). These functions were used to transform variables at their optimal scales in univariate GLMs, as described above. Unmodified linear forms were also included, resulting

Table 2 Summary of model performance statistics for single-scale optimized (SSO) models, single-scale shape-optimized (SSSO) models, the multi-scale optimized (MSO) model, and the multi-scale and shape-optimized (MSSO) model

Model	Threshold	PCC	Sensitivity	Specificity	Kappa	AUC
SSO—250 m	0.37	0.8700	0.5965	0.9462	0.5873	0.85
SSSO—250 m	0.46	0.8719	0.5175	0.9707	0.5651	0.86
SSO—500 m	0.41	0.8719	0.6140	0.9438	0.5975	0.87
SSSO—500 m	0.41	0.8681	0.5877	0.9462	0.5798	0.88
SSO—1 km	0.39	0.8413	0.5526	0.9218	0.5047	0.86
SSSO—1 km	0.35	0.8375	0.6579	0.8875	0.5336	0.87
SSO—2 km	0.27	0.8145	0.8246	0.8117	0.5391	0.88
SSSO—2 km	0.27	0.8222	0.7632	0.8386	0.5357	0.88
SSO—4 km	0.32	0.8509	0.7456	0.8802	0.5886	0.90
SSSO—4 km	0.33	0.8547	0.7807	0.8753	0.6062	0.90
SSO—8 km	0.40	0.8700	0.6667	0.9267	0.6087	0.91
SSSO—8 km	0.45	0.8795	0.7018	0.9291	0.6410	0.91
SSO—16 km	0.54	0.8929	0.6491	0.9609	0.6601	0.93
SSSO—16 km	0.54	0.8891	0.6667	0.9511	0.6551	0.93
MSO	0.37	0.8891	0.7456	0.9291	0.6747	0.92
MSSO	0.43	0.8815	0.7018	0.9315	0.6455	0.93

in five functional forms tested in this model approach. The best performing shape for each variable was determined by lowest AICc value before conducting, as previously described, filtering based on *p*-values and collinearity. In the final multivariate GLM for each scale, for quadratic variables, we included both linear and quadratic terms as separate explanatory variables. We were able to include these quadratic terms using ‘biostats’ (McGarigal 2018), while running all combinations of models, providing an advantage over other similar approaches, such as *dredge* (Bartoń 2018).

Model 4: Multi-scale- and shape-optimized model approach

In the multi-scale and shape-optimized (MSSO) model, we include both scale- and shape-optimization. As in MSO and SSSO models, we determined the best performing shape and scale for each variable by lowest AICc value before conducting filtering based on *p*-values and collinearity. In effect, variables were included in a final multivariate GLM at their optimal scale and functional form. As in the SSSO approach, for quadratic variables, we included both linear and quadratic terms as separate explanatory variables using ‘biostats’ (McGarigal 2018).

Evaluating model performance and variable importance

We evaluated model performance in three ways. First, we calculated sensitivity (*correctly predicted presence/total points present*), specificity (*correctly predicted absence/total points absent*), percent correctly classified (*PCC; summary of correctly predicted points at a defined threshold*), Kappa (*proportion correctly predicted points accounting for probability of chance*), and area under the ROC curve (AUC). We measured variable importance via AICc variable importance from the model averaging table and magnitude of the standardized coefficients (Wasserman et al. 2012). We also tested the influence of each variable importance from the model averaging table and magnitude of the standardized coefficients (Wasserman et al. 2012). We also tested the influence of each variable on model performance by calculating the difference in probability of tiger detections when the variable increases from the 10th to 100th

percentile, using standardized values and holding all other variables constant at their medians. Variables were back-transformed to original values and plotted to evaluate changes in probability of tiger presence relative to changes in each variable. Lastly, we generated prediction maps to investigate differences in predictions for each modelling approach. These maps depict spatial variation in predicted tiger presence, with each cell containing a predicted presence value ranging from 0 (absent) to 1 (present). To compare predictive maps between the four top models, we used three approaches. First, we calculated the correlation between the pixel values in each of the four maps. Second, we calculated the average absolute difference between the pixel values in the four maps. Third, we classified the four maps into high quality (greater than probability of 0.7) and mid-high quality habitat (greater than probability of 0.5). We then used FRAGSTATS (McGarigal et al. 2012) to calculate percentage of landscape (PLAND), patch density (PD), correlation length (GYR) and largest patch index (LPI) for each of the maps.

Results

Model performance

All models demonstrated high predictive performance, with AUC ranging from a low of 0.85 (for the 250 m single-scale optimized [SSO] model) to 0.93 for the multi-scale and shape-optimized (MSSO) model and the top single scale-optimized 16 km models (SSO and SSSO; Table 2). Similarly, the Kappa statistic, percent observations correctly classified, model sensitivity, and specificity were high for all models, with the 16 km SSO, 16 km SSSO, MSO, MSSO models generally out-performing smaller-scale SSO models. There was slight disagreement among these measures in terms of which of these models had the highest performance. The 16 km SSO, 16 km SSSO, and MSSO models had the highest performance based on AUC (0.93), MSO had the highest performance based on Kappa, and the 16 km SSO showing the highest performance based on PCC. All four of these models had very high performance, with AUC over 0.90, PCC greater than 0.88, and Kappa greater than 0.64, which demonstrates exceptionally high

ability to correctly discriminate between detection and non-detection points.

Model interpretation

Although the top models (16 km SSO, 16 km SSSO, MSO, and MSSO) each had very high and very similar predictive performance, they differed substantially in model interpretation. We base model interpretation on several characteristics: (1) variables included in the model, (2) scale at which each variable is included, (3) sign of the variables included, (4) variable importance, (5) the shape of variables included in SSSO and MSSO models, and (6) differences in prediction maps.

Variables included

Camera effort (Cam_eff) was included in all four of the top models (SSO, MSO, SSSO, and MSSO), showing the universally high importance of sampling effort to model prediction. Other than camera effort, only percentage of bamboo forest (FT5_PLAND) was included in all four top models (with a positive coefficient), indicating that the extent of bamboo forest surrounding a sampling location is a ubiquitously important predictor of tiger occurrence in the study area. No other variable was included in all four of the top models, although secondary growth forest (FT_6) was included in all four top models based on slightly different FRAGSTATS metrics, which measure extensiveness of the cover type. Specifically, correlation length of secondary growth forest (FT6_GYR) was included in MSO, the 16 km SSSO, and MSSO, while percentage of secondary growth forest (FT6_PLAND) was included in the 16 km SSO. This suggests that the extensiveness of secondary growth forest is also an important predictor of tiger occurrence in the study area (with a negative coefficient). Similarly, the standard deviation of compound topographic index (CTI_SD), a measure of topographical complexity, was included in MSSO and the 16 km SSO. Focal mean of topographical roughness index (TRI_FM) and standard deviation of elevation (DEM_SD), which measure similar attributes of topographical complexity, were also included in MSSO while standard deviation of elevation was also included in the 16 km SSSO. This suggests that topographical heterogeneity is a consistently important predictor of tiger occurrence in the study area.

Several variables were included in two of the four top models, with focal extensiveness of open forest (20–40% forest cover; TC2_PLAND), included in both MSO and MSSO, but not in the top SSO and SSSO models. Similarly, focal extensiveness of shrubland/grassland (LC4_PLAND), was included in the MSO and 16 km SSO/SSSO. Four variables were included in only one of the four top models, with percentage evergreen forest (FT1_PLAND) and percentage mosaic cropland (LC2_PLAND) included only in the 16 km SSSO, correlation length of reforested areas (FT4_GYR) included only in MSSO, and correlation length of urban areas (LC6_GYR) included only in the 16 km SSO. Three variables were shared between the 16 km SSO and SSSO models while four variables were shared between both shape-optimized models (SSSO and MSSO).

Variable scales

Variables were generally represented at broad spatial scales. Almost all variables in the MSSO (six out of seven; Table 3) and MSO (five out of seven; Online Resource 2, Table 4) were selected at the broadest spatial scales (8 and 16 km). In MSO, the smallest scale represented was 1 km, while the smallest scale represented in the MSSO was 4 km. This representation of broad scales is also evident in the similarly strong predictive performance of both single-scale optimized (i.e. SSO and SSSO) 16 km models compared with multi-scale optimized models. It is notable that the worst performing model was the SSO developed at the finest spatial scale (250 m).

Variable sign

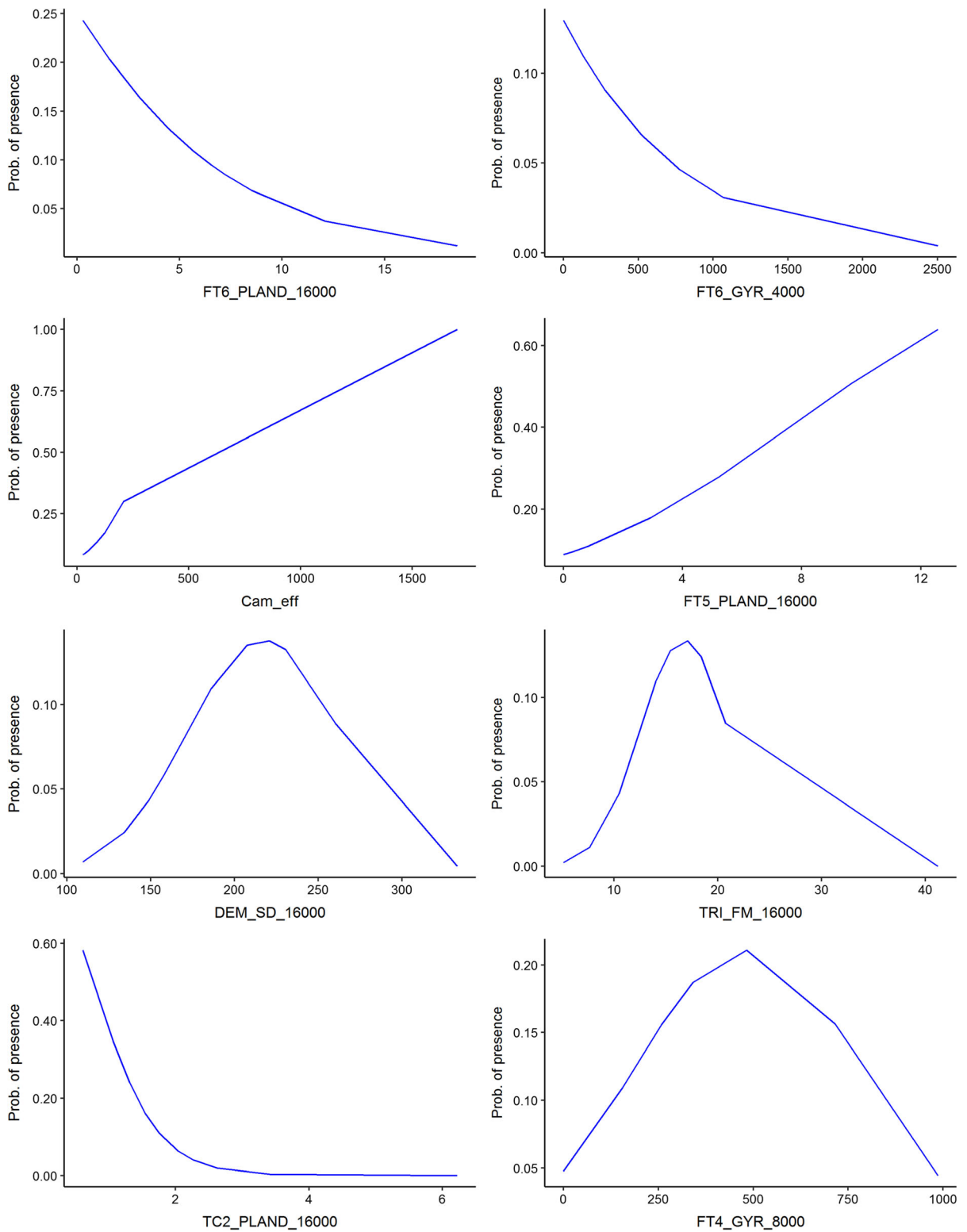
The signs of shared variables remained consistent across our four top models. Variables with a positive relationship with tiger occurrence included camera effort (Cam_eff) and percentage of bamboo forest (FT5_PLAND) in all models, standard deviation of compound topographic index (CTI_SD) in the 16 km SSO and MSO models, standard deviation of terrain roughness index (TRI_SD) in the MSSO model, and percentage of evergreen forest (FT1_PLAND) in the 16 km SSSO. Variables with a negative relationship included correlation length of secondary forest (FT6_GYR) in the 16 km SSSO, MSO, and MSSO, percentage of open forest (TC2_PLAND) in both

Table 3 Generalized linear model results for the multi-scale and shape-optimized (MSSO) model, predicting tiger presence based on 1166 detections of tiger in DPKY (2008–2017),including standardized regression coefficients (β), adjusted standard error, z-score (z), and significance (p)

Variable	ID	β	SE (Adj)	z	p
(Intercept)	–	– 1.367	0.314	4.352	< 0.001
% Secondary Forest (16 km) ^{LOG}	FT6_PLAND_16000_LOG	– 0.777	0.210	3.702	< 0.001
Correlation length Secondary Forest (4 km) ^L	FT6_GYR_4000L	– 0.701	0.235	2.978	0.003
Camera effort (# trap nights)	Cam_eff	1.460	0.309	4.718	< 0.001
% Bamboo (16 km) ^L	FT5_PLAND_16000	1.028	0.525	1.960	0.050
DEM Standard Deviation (16 km) ^L	DEM_SD_16000	0.995	0.308	3.226	0.001
% Bamboo (16 km) ^Q	FT5_PLAND_16000Q	– 0.072	0.232	0.309	0.757
DEM Standard Deviation (16 km) ^Q	DEM_SD_16000Q	– 0.866	0.336	2.575	0.010
Focal mean of terrain roughness index (16 km) ^L	TRI_FM_16000L	1.165	0.502	2.317	0.020
Focal mean of terrain roughness index (16 km) ^Q	TRI_FM_16000Q	– 1.671	0.579	2.887	0.004
% Open forest (16 km) ^L	TC2_PLAND_16000_L	– 2.353	0.457	5.149	< 0.001
Correlation length of reforested areas (8 km) ^L	FT4_GYR_8000_L	0.968	0.321	3.011	0.003
Correlation length of reforested areas (8 km) ^Q	FT4_GYR_8000_Q	– 0.555	0.195	2.844	0.004

Shape (i.e. functional form) of each variable is indicated by: ^Llinear, ^QQuadratic, ^{LOG}Logarithmic**Table 4** Summary of variable importance among top models, the single-scale optimized (SSO) 16 km model, the single-scale shape-optimized (SSSO) 16 km model, the multi-scale optimized (MSO) model and the multi-scale and shape-optimized (MSSO) model. Specific metrics include AICcvariable importance, standardized coefficients, and change in probability of tiger presence with change in variable value from 10th to 100th percentiles. The SSSO and MSSO models included both linear (^L) and quadratic (^Q) terms for five variables

	AICc variable importance				Standardized coefficients				Δ Probability of tiger presence			
	SSO	SSSO	MSO	MSSO	SSO	SSSO	MSO	MSSO	SSO	SSSO	MSO	MSSO
Cam_eff	2	5	6	3	1.355	1.487	1.415	1.460	0.97	0.94	0.96	0.92
CTI_SD	1	–	3	–	0.720	–	0.616	–	0.22	–	0.44	–
DEM_SD	–	4	–	1	–	0.333 ^L – 0.868 ^Q	–	0.995 ^L – 0.866 ^Q	–	0.07	–	0.13
FT1_PLAND	–	1	–	–	–	0.733	–	–	–	0.12	–	–
FT4_GYR	–	–	–	1	–	–	–	0.968 ^L – 0.555 ^Q	–	–	–	0.17
FT5_PLAND	2	5	6	2	1.964	1.556	0.988	1.028 ^L – 0.072 ^Q	0.90	0.70	0.46	0.55
FT6_GYR	–	2	5	3	–	– 0.443	– 0.675	– 0.701	–	– 0.10	– 0.06	– 0.13
FT6_PLAND	2	–	–	2	– 1.827	–	–	– 0.777	– 0.11	–	–	– 0.23
LC2_PLAND	–	5	–	–	–	– 1.130 ^L – 0.770 ^Q	–	–	–	– 0.10	–	–
LC4_PLAND	2	2	1	–	– 0.647	– 0.496	– 0.486	–	– 0.08	– 0.13	– 0.06	–
LC6_GYR	1	–	–	–	– 0.613	–	–	–	– 0.04	–	–	–
TC2_GYR	–	–	2	–	–	–	– 0.704	–	–	–	– 0.11	–
TC2_PLAND	–	–	5	1	–	–	– 1.437	– 2.353	–	–	– 0.02	– 0.58
TRI_FM	–	–	–	2	–	–	–	1.165 ^L – 1.671 ^Q	–	–	–	0.13
TRI_SD	–	–	2	–	–	–	0.384	–	–	–	0.38	–



◀ **Fig. 2** Partial effects plots for the multi-scale and shape-optimized (MSSO) model showing changes in predicted probability of tiger presence when variables increase from their 10th to 100th percentile, holding other variables at their medians. Graphic generated in R (R Development Core Team 2017)

MSSO and MSSO, correlation length of open forest (TC2_GYR) in MSSO, percentage of shrubland/grassland (LC4_PLAND) in the 16 km SSO, 16 km MSSO, and MSSO models, and correlation length of urban areas (LC6_GYR) in the 16 km SSO model. Three variables in shape-optimized models (16 km MSSO and MSSO) exhibited a positive parabolic relationship between probability of tiger presence and increasing variable values (Fig. 2; Online Resource 3, Fig. 4). These included standard deviation of elevation (DEM_SD; MSSO/MSSO), focal mean of terrain roughness index (TRI_FM; MSSO), and correlation length of reforested areas (FT4_GYR; MSSO). This indicates probability of tiger presence peaks at intermediate values of these variables.

Variable importance

In the 16 km SSO model, AICc variable importance was similar across variables with the number of model subsets in which variables were represented ranging from $n = 1$ to 2 (Table 4). Key variables in this model, as determined by standardized coefficients, included percentage of bamboo forest (FT5_PLAND; $\beta = 1.964 \pm 0.314$), percentage of secondary forest (FT6_PLAND; $\beta = -1.827 \pm 0.424$), and camera effort (Cam_eff; $\beta = 1.355 \pm 0.279$). Variables which demonstrated the strongest effect on probability of tiger presence were camera effort ($\Delta 0.97$), percentage of bamboo forest ($\Delta 0.90$) and standard deviation of compound topographic index ($\Delta 0.22$).

The 16 km MSSO model exhibited different patterns in AICc variable importance, with the number of model subsets in which variables were represented ranging from $n = 1$ to 5. Variables with notable standardized coefficients, included percentage of bamboo forest (FT5_PLAND; $\beta = 1.556 \pm 0.440$), percentage of mosaic cropland (LC2_PLAND; $\beta = -1.130 \pm 0.495$; $Q \beta = -0.770 \pm 0.516$), and camera effort (Cam_eff; $\beta = 1.487 \pm 0.311$). The strongest effect on probability of tiger presence were

evident in camera effort ($\Delta 0.94$) and percentage of bamboo forest ($\Delta 0.70$).

The MSSO model had notable variation in AICc variable importance, ranging from $n = 1$ for percentage shrubland/grassland (LC4_PLAND) to $n = 6$ for camera effort (Cam_eff) and percentage of bamboo forest (FT5_PLAND). Variable importance, in terms of the magnitude of standardized coefficients, was strongest for percentage open forest (TC2_PLAND; $\beta = -1.437 \pm 0.349$), camera effort (Cam_eff; $\beta = 1.415 \pm 0.299$), and percentage of bamboo forest (FT5_PLAND; $\beta = 0.988 \pm 0.238$). Changes in probability of tiger presence were greatest among camera effort ($\Delta 0.96$), standard deviation of compound topographic index (CTI_SD; $\Delta 0.46$), percentage of bamboo forest ($\Delta 0.44$), and standard deviation of terrain roughness index (TRI_SD; $\Delta 0.38$).

The MSSO model had some degree of variation in AICc variable importance, ranging from $n = 1$ for standard deviation of elevation (DEM_SD), correlation length of reforested areas (FT4_GYR), and percentage of open forest (TC2_PLAND) to $n = 3$ for camera effort (Cam_eff) and correlation length of secondary forest (FT6_GYR). Standardized coefficients indicated the importance of percentage of open forest (TC2_PLAND; $\beta = -2.353 \pm 0.457$), camera effort ($\beta = 1.460 \pm 0.309$), focal mean of terrain roughness index (TRI_FM; $\beta = 1.165 \pm 0.502$), $\beta = -1.671 \pm 0.579$ and percentage of bamboo forest (FT5_PLAND; $\beta = 1.028 \pm 0.525$), $\beta = -0.072 \pm 0.232$). Variables with the strongest effect on probability of tiger presence included camera effort (Cam_eff; $\Delta 0.92$), percentage of open forest (TC2_PLAND; $\Delta -0.58$), and percentage of bamboo forest ($\Delta 0.55$).

Shape of variables

While five functional forms were tested during variable selection for the 16 km MSSO and MSSO models—linear, quadratic, logarithmic, exponential, and negative exponential—only three were represented in fully averaged models. Quadratic relationships (for which linear terms were also included) were evident in percentage of bamboo forest (FT5_PLAND) at 16 km (MSSO/MSSO), standard deviation of elevation (DEM_SD) at 16 km (MSSO/MSSO), focal mean of terrain roughness index (TRI_FM) at 16 km (MSSO), percentage of mosaic

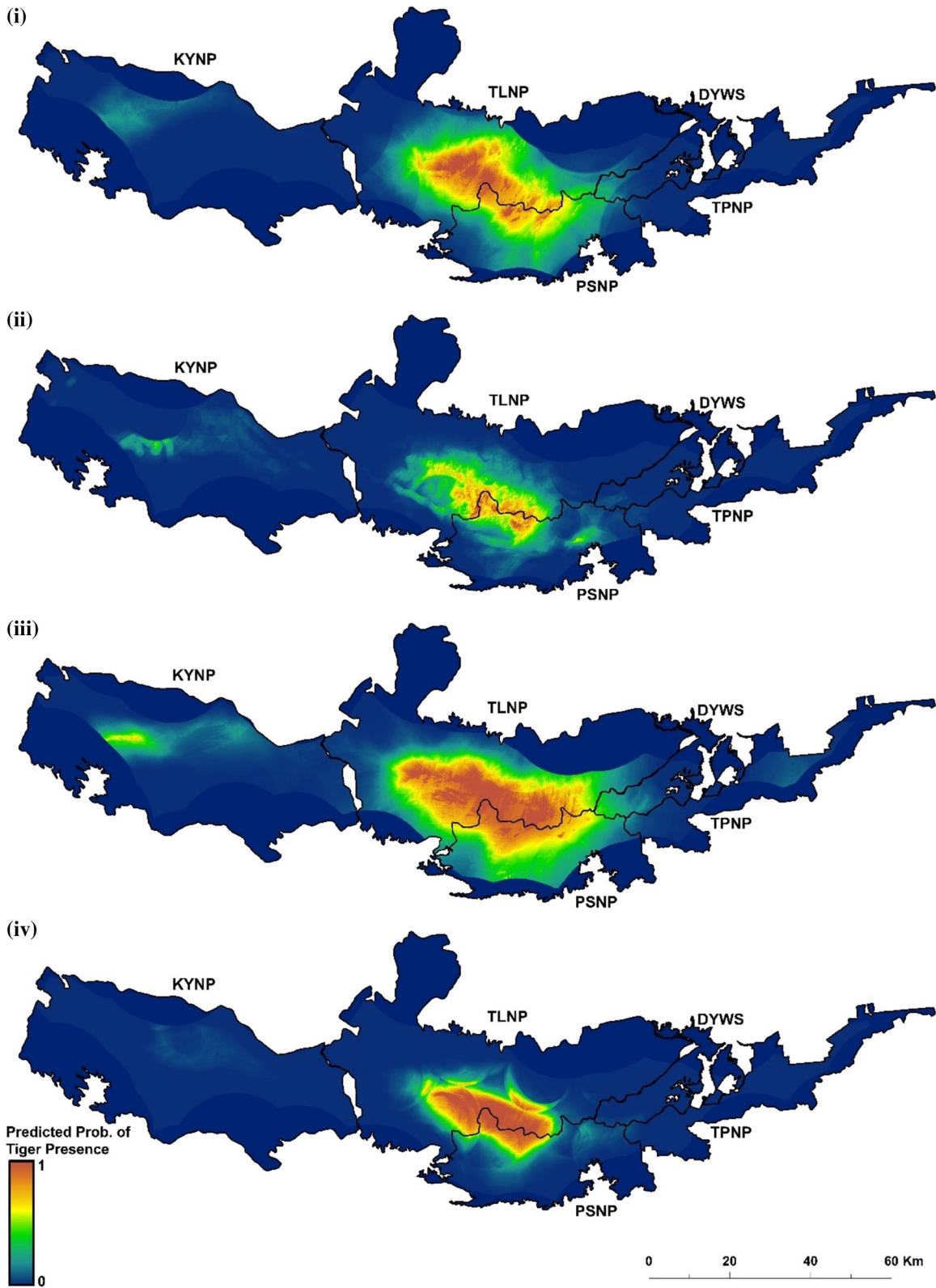


Fig. 3 Prediction maps of probability of tiger presence, with values ranging from 0 (predicted absent) to 1 (predicted present), generated from the top four fully averaged models—(i) the single-scale optimized (SSO) 16 km model, (ii) the scale-optimized (MSO) model, the single-scale shape-optimized (SSSO) 16 km model, and (iv) the scale- and shape-optimized (MSSO) model. Map generated with ArcGIS 10.3.1 (Environmental Systems research Incorporated, ESRI, Redlands, CA, USA, 2011)

cropland (LC2_PLAND) at 16 km (SSSO) and correlation length of reforested areas (FT4_GYR) at 8 km (MSSO). Three variables—percentage of secondary forest (FT6_PLAND) at 16 km (MSSO), correlation length of secondary forest (FT6_GYR) at 16 km (SSSO), and percentage of evergreen forest (FT1_PLAND) at 16 km (SSSO)—exhibited logarithmic relationships. Linear relationships were observed in correlation length of secondary forest (FT6_GYR) at 4 km (MSSO), percentage of reforested areas (LC4_PLAND; SSSO), and percentage of open forest (TC2_PLAND; MSSO) at 16 km.

Model prediction

Correlation between prediction maps (Fig. 3) for the four top models was reasonably high. Correlation of the prediction maps for SSSO–SSO was the highest (0.95), followed by MSSO–MSO (0.92), and MSSO–SSO (0.84; Table 5).

Differences between pixel values were generally highest in the central part of the landscape, within TLNP and PSNP, with SSSO and SSO 16 km models having a notably higher predicted presence of tigers over a broader area compared to MSSO and MSO models (Online Resource 3, Fig. 5). Between these models, SSSO had higher predicted values elsewhere in the complex. Predicted tiger presence was higher in a smaller core area of the complex in MSSO and lower predicted presence outside this area compared with other models. The MSO generally had lower predicted tiger presence in more central areas compared to other models though predicted higher probability of presence in certain areas elsewhere in the complex such as in small areas of PSNP and KYNP. Average absolute difference was highest between SSSO and MSO maps (0.17) and lowest between MSSO and MSO maps (0.04; Table 4).

Percentage of high-quality habitat (> 0.7 predicted tiger presence) in the landscape was highest in the 16 km SSSO (~ 12.3%), followed by the MSSO

Table 5 Matrix of correlation (COR) and average absolute difference (AAD) between prediction maps between the single-scale optimized (SSO) 16 km model, multi-scale optimized (MSO) model, single-scale shape-optimized (SSSO) 16 km model, and multi-scale and shape-optimized (MSSO) model

	SSO	MSO	SSSO	MSSO	
SSO		0.09	0.10	0.08	<i>AAD</i>
MSO	0.84		0.17	0.04	
SSSO	0.95	0.83		0.15	
MSSO	0.84	0.92	0.83		
		<i>COR</i>			

(~ 5.5%), the 16 km SSO (~ 4.7%), and was notably low in MSO (~ 0.7%; Table 6). Extent of mid-high quality habitat (> 0.5 predicted tiger presence) was also highest in the 16 km SSSO (~ 20.8%), followed by the 16 km SSO (~ 11.7%), MSSO (~ 7.9%), and MSO (~ 4.1%). Mid- to high-quality habitat was generally more dispersed in MSO than other models, as indicated by relatively higher PD, and lower LPI and GYR values, characterized by several distinct patches. In the 16 km SSSO/SSO and MSSO, both mid-high and high quality habitat were largely represented by single patches.

Discussion

The focus of this study was to quantitatively evaluate the relative impacts of scale- and functional shape-optimization on model performance, prediction and interpretation in habitat selection studies, using a globally important tiger population as a case study. Our study confirms the scale-dependence of habitat selection in tigers, and our results suggest it is important for further confirm the importance of habitat selection research to formally evaluate and optimize scale as a component of the modelling process. In contrast, we did not find strong and clear effects of optimizing models to account for different, nonlinear functional shapes of relationship between species occurrence and environmental variables.

Consistent with our first hypothesis, we found that tiger occurrence probability is highly scale dependent. There were strong differences in the univariate strength of relationship between most variables across scale. In addition, we found large differences in model

Table 6 Landscape metrics of high (> 0.7) and mid-high (> 0.5) probability of tiger presence among the four top models. Landscape metrics include percentage of landscape

(PLAND), patch density (PD), largest patch index (LPI), and radius of gyration/correlation length (GYR)

Habitat Quality	Model	PLAND	PD	LPI	GYR
HIGH	SSO	4.6680	0.0005	4.668	6106
	MSO	0.7400	0.0013	0.275	1557
	SSSO	12.3025	0.0003	12.303	9767
	MSSO	5.4998	0.0005	5.388	6812
MID-HIGH	SSO	11.6668	0.0003	11.6668	9599
	MSO	4.0950	0.0059	4.0534	6673
	SSSO	20.7904	0.0006	20.7506	12,173
	MSSO	7.8861	0.0008	7.7687	7739

performance and interpretation, indicating that the ability to predict tiger occurrence is strongly affected by the scale in which variables are measured. This scale-dependence in model interpretation is particularly stark between top models and smaller scale single-scale models (SSO/SSSO). Nine variables in top models were not present in fine-scale single-scale optimized (SSO) models (250 m–4 km) and ten variables represented in these fine-scale SSOs/SSSOs were not present in top models. This is particularly important, as most past tiger habitat modelling, similar to past modelling for other large carnivores (Mateo-Sánchez et al. 2015), have used SSO models with a relatively fine (~ 500 m–1 km) focal scale.

Our results reinforce that the lens with which the niche of organism is evaluated, specifically its relationships with habitat, has a considerable influence in the interpretation of models. This high degree of scale-dependence is evident in other tiger studies (Rostro-García et al. 2016; Krishnamurthy et al. 2016; Reddy et al. 2017), as well as studies on cheetah (Rostro-García et al. 2015), leopard (Pitman et al. 2017; Kittle et al. 2018), puma (Zeller et al. 2014) and clouded leopard (Hearn et al. 2018; Macdonald et al. 2019). Substantial differences in model interpretation and identification of optimal habitat based on the scale of analysis is also consistent with numerous studies of other taxonomic groups (Thompson and McGarigal 2002; Wasserman et al. 2012; Mateo-Sánchez et al. 2014; Wan et al. 2017). Scale-dependence in habitat relationships is a foundational idea in wildlife ecology (Wiens 1989; Levin 1992; Thompson and McGarigal 2002) and critical for analyses which aim to quantify

species niches (Hegel et al. 2010). Our results strongly reinforce this and demonstrate the importance of the scale of analysis for predicting tiger presence.

A majority of variables were most strongly associated with tiger occurrence at the broadest spatial scales, which was consistent with our second hypothesis. In the multi-scale optimized (MSO) model, 71% of the included variables were measured at the broadest (16 km) or second broadest (8 km) scale. Similarly, in the multi-scale and shape optimized (MSSO) model, 88% were selected at the broadest or second broadest scale, with more than half at the broadest scale. Notably, there were clear patterns of improvement in AUC in both single-scale model approaches (SSO/SSSO) from fine to broad scales. The 16 km SSO and SSSO models, in which all variables were measured at the broadest scale, were the best performing among these single-scale models. The high consistency among these top models clearly suggests that habitat selection for tigers in our study is largely related to broad-scale habitat factors. These results are aligned with previous studies on tigers (Krishnamurthy et al. 2016; Rostro-García et al. 2016; Reddy et al. 2017), brown bear (Mateo-Sánchez et al. 2015), leopard (Kittle et al. 2018), and clouded leopard (Hearn et al. 2018), all of which found dominant effects of environmental variables at broad scales. Importantly, however, most habitat relationship studies, even for large bodied and highly mobile animals, typically have utilized single-scale analysis of fine scale (e.g. 500 m–1 km) environmental data (McGarigal et al. 2016). Our analysis suggests that such fine-scale analysis would fail to accurately or

fully elucidate species habitat selection patterns, even if model performance is reasonably strong, such as with single, fine-scale models in our study.

Our third hypothesis was that the multi-scale optimized models (MSO and MSSO) would outperform any single-scale optimized (SSO/SSSO) model. This follows from arguments and observations made by McGarigal et al. (2016) who reviewed a large number of habitat modelling papers and found that, in nearly all cases that were formally evaluated, multi-scale optimization outperformed single-scale modelling of habitat relationships. The habitat niche of any species consists of a number of dimensions, each composed of environmental variation at particular characteristic scales. Our results are not fully consistent with this third hypothesis. Both the multi-scale optimized and the best single-scale optimized models (16 km SSO/SSSO) performed exceptionally and equivalently well. These results suggest that a single-scale model may sometimes perform as well as a multi-scale optimized model, but only when organism-habitat relationships are dominated by a single-scale response and that the appropriate scale of analysis is evaluated. This has also been observed in other studies (Martin and Fahrig 2012; Krishnamurthy et al. 2016; Timm et al. 2016). In the case of tigers (Krishnamurthy et al. 2016; Reddy et al. 2017), a single, broad-scale model might produce very similar results to a multi-scale optimized model. However, we emphasize that it is impossible to assess the appropriate scale of analyses for a single-scale model a priori. Further, even when a single, large scale may be optimal, broad application can result in potentially important fine-scale niche dynamics being under-represented (Mateo-Sanchez et al. 2014) leading to oversimplified and potentially misleading interpretations of model results. Therefore, even when a single-scale model provides the strongest prediction, it requires use of scale-optimization to identify and confirm this. Further, when a single-scale model isn't optimal, multi-scale optimization is required to identify the optimal multi-scale model.

Perhaps the most novel aspect of our study is the comparison of the multi-scale and shape-optimized model (MSSO) with multi-scale optimized (MSO) and single-scale optimized (SSO/SSSO) models. We expected that the multi-scale and shape-optimized model and single-scale shape-optimized (SSSO) models would out-perform non-shape-optimized models.

This is because, like scale-dependence, species occurrence probability may have a different functional response shape for each variable, with some responses being linear, some unimodal, and some curvilinear (Austin et al. 1990). Results of similar studies suggest that optimizing for functional form may have meaningful influence in the selection of top models (DeVoe et al. 2015; Mateo-Sánchez et al. 2015), potentially allowing for the expression of otherwise obscure environmental effects (Carle 2006), which may improve explanatory power. We hypothesised that optimizing the response shape of variables, like optimizing the scale of response, would measurably improve the predictive ability of the model. Our results do not support this.

While the multi-scale and shape-optimized models were tied as the highest performing model based on the most widely used criterion (AUC), the 16 km single-scale optimized model performed equally well. Furthermore, the multi-scale optimized and the single-scale optimized 16 km models performed slightly better based on other measures. The single-scale optimized (SSO) and single-scale shape-optimized (SSSO) models exhibited similar patterns in model performance, with AUC values generally increasing with scale. While AUC was slightly higher (0.01) in finer-scale shape-optimized models compared to corresponding non-optimized models, this was not the case at broader-scales. This suggests that shape-optimization may be less important than scale-optimization.

While model performance was not substantially improved by shape-optimization in our study, the inclusion of variables at their optimal functional form clarified relationships between habitat characteristics and species response. For example, topographic variables (i.e. elevation and terrain roughness) are present exclusively in non-linear forms (i.e. quadratic and logarithmic) in shape-optimized models (SSSO/MSSO). In effect, the multi-scale and shape-optimized model suggested a more complex relationship between species presence and topographic heterogeneity than in non-optimized models. Expression of these types of relationships are also evident in other species-habitat modelling studies, in particular, the representation of quadratic relationships in final models (Mashintonio et al. 2014; DeVoe et al. 2015; Timm et al. 2016; Macdonald et al. 2018; Macdonald et al. 2019). Notably, inclusion of variables at their

optimal functional form influenced representation of variables in final, fully averaged models with differences in variables included between optimized (SSSO) and non-optimized (SSO) models. For example, in 8 and 16 km models, only two variables were included in both SSSO and SSO models out of 12 and 11 total variables in these models, respectively. Further, only two out of 11 total variables were represented in both SSO and MSSO models. This illustrates that models may be sensitive to optimizing for functional form, as with scale, and may lead to considerable differences in model interpretation and conclusions about species–habitat relationships.

Limitations/caveats

Our fundamental goal in this study was to evaluate the effect of scale- and shape-optimization on the performance, prediction and interpretation of habitat relationships models. We use a single case study species and system (tigers in Thailand) for this evaluation, which limits the generality of our conclusions, however, in exchange we feel our example provides realism in that it reflects actual relationships of a focal species of ecological and conservation importance. We suggest future work employ simulation studies which can control scale and shape effects and, therefore, more reliably and completely evaluate the sensitivity of modelling methods to scale- and shape-optimization. Second, given our focus was on exploring the novel question of impacts of scale- and shape-optimization simultaneously, rather than evaluating the best approaches for scale optimization, our example uses GLM modelling because it is the most well understood and widely-used method in habitat modelling, and the dominant approach in multi-scale habitat modelling studies published to date (McGarigal et al. 2016). This approach to scale selection is consistent with many other species-habitat modelling studies (Rostro-García et al. 2016; Kittle et al. 2018; Macdonald et al. 2018). However, with the rapid emergence and adoption of other modelling methods, in particular machine learning approaches (e.g., Evans et al. 2011; Cushman and Wasserman 2018), we suggest future work should focus on comparing performance of scale- and shape-optimization in a wider range of modelling approaches. In particular, tree-based machine learning approaches like random forests perhaps can automatically optimize nonlinear

functional shape (Evans et al. 2011), which could be a great advantage, while others, like Maxent, do not (Elith et al. 2011). While there are a myriad of approaches for habitat–selection analysis, we opted to utilize GLMs given their intuitiveness, flexibility, and their wide-spread use in studies investigating scale in similar studies (McGarigal et al. 2016).

We also note that our definition of optimization closely follows the definition proposed by McGarigal et al. (2016) of pseudo-optimization as the scales we evaluated were selected *a priori*. Importantly, this approach does not optimize scale relationships across a continuous range, and does not optimize them multivariately and simultaneously. A true optimized approach may have allowed for inclusion of more specific scales that best explain variation in tiger occurrence. However, we feel that the scales selected for evaluation were sufficient for understanding relevant scales of relationships between covariates and tiger occurrence and, most importantly, allowed for a more effective and straightforward evaluation of scale-dependence in our study compared to a true-optimized approach. However, improvement of scale-optimization approaches is an important current topic in habitat relationships modelling (McGarigal et al. 2016), and we strongly suggest future research to explore the performance of a range of continuous and multivariate simultaneous optimization approaches.

Species management implications

We recommend central DPKY, particularly areas of closed forest cover containing bamboo forest patches within a 16 km window, should be managed as an area of high conservation priority for tigers in this landscape. Our models reinforce the importance of protecting surrounding broad-scale (8 to 16 km) forest within protected areas with low human impact as part of a landscape-scale management strategy. Protection of this core area and facilitating unconstrained movement to other available habitat will be critical to the long-term recovery of this population. Scale-dependence in species-habitat relationships, as demonstrated by this and other studies (Wasserman et al. 2012; Mateo-Sanchez et al. 2014) can have monumental implications for species management. As such, efforts to develop management and recovery strategies should account for scale-dependence of tigers in this landscape. Our results are consistent with other

literature highlighting the importance of large protected areas and low human disturbance for tiger populations (Trisurat et al. 2010; Sunarto et al. 2012; Ngoprasert et al. 2012; Hebblewhite et al. 2014; Reddy et al. 2017).

It is pertinent to highlight that our results reflect an expression of the tigers' realized niche within DPKY that may be severely constrained in comparison with its theoretical fundamental niche. The tigers' range itself has suffered considerable declines and its current range likely reflect its refuge in areas less likely to be impacted by human activities, which are the chief driver of range and population declines (Goodrich et al. 2015). In our study area, areas of the landscape with flat topography are typically highly affected by human activities, which may explain their lower selection. Historically, Thailand's lowland forests, now converted by human activity, would have likely represented preferable habitat due to high prey densities (Rabinowitz 1993; Sunquist et al. 1999).

Conclusion

While shape-optimization did not substantially improve performance over other models, it did allow for the expression of potentially important relationships between tigers and covariates that were not apparent in the models assuming linear forms. Importantly, our study clearly reinforces previous studies (McGarigal and Cushman 2002; McGarigal et al. 2016) which highlight the importance of accounting for scale when modelling habitat selection and other ecological relationships, particularly for wide-ranging species such as tigers. We recommend that habitat-selection studies on tigers and other species incorporate a robust, ecologically-relevant scale-optimization framework and consider the inclusion and evaluation of shape-optimization in model development.

Acknowledgements We would like to extend our deepest thanks to Thailand's Department of National Parks, Wildlife and Plant Conservation (DNP), particularly Dr. Somphot Duangchantrasiri, Dr. Saksit Simcharoen, and the DNP's protected area rangers. Data were generated from tiger survey work conducted by the DNP and Freeland Foundation, supported in part by Panthera and WWF-Thailand. Funding support was provided by the U.S. Fish and Wildlife Service Rhinoceros and Tiger Conservation Fund, David Shepherd Wildlife Foundation, Care for the Wild International/Born Free Foundation, 21st Century Tiger, and Point Defiance Zoo &

Aquarium. Authors EA and ZK were supported by the Robertson Foundation via grants to DWM. We thank Dr. Kevin McGarigal for assistance during analysis. We also wish to thank Dr. Songtam Suksawang, park managers and staff, Christopher Hallam, Luke Stokes, Paul Thompson, and Dr. Rob Steinmetz.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ash E, Kaszta Ž, Noochdumrong A, Redford T, Chanteap P, Hallam C, Jaroensuk B, Raksat S, Srinoppawan K, MacDonald DW (2020) Opportunity for Thailand's forgotten tigers: assessment of Indochinese tiger *Panthera tigris corbetti* and prey from camera-trap surveys in Eastern Thailand. *Oryx*, 1–8. <https://doi.org/10.1017/S0030605319000589>
- Austin MP, Nicholls AO, Margules CR (1990) Measurement of the realized qualitative niche: environmental niches of five Eucalyptus species. *Ecol Monogr* 60:161–177
- Bar-Massada A, Wood EM, Pidgeon AM, Radeloff VC (2011) Complex effects of scale on the relationships of landscape pattern versus avian species richness and community structure in a woodland savanna mosaic. *Ecography* 35:393–411
- Barber-Meyer SM, Jnawali SR, Karki JB, Khanal P, Lohani S, Long B, Mackenzie DI, Pandav B, Pradhan NMB, Shrestha R, Subedi N, Thapa G, Thapa K, Wikramanayake E (2013) Influence of prey depletion and human disturbance on tiger occupancy in Nepal. *J Zool* 289:10–18
- Bartoń K (2018) Multi-Model Inference (MuMIn). Comprehensive R Archive Network (CRAN), Version 1.40.4
- Bates D, Mächler M, Bolker B, Walker S (2014) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48
- Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. *Hydrol Sci Bull* 24:43–69
- Blonder B (2018) Hypervolume concepts in niche- and trait-based ecology. *Ecography* 41:1441–1455
- Bosco L, Wan HY, Cushman SA, Arlettaz R, Jacot A (2019) Separating the effects of habitat amount and fragmentation on invertebrate abundance using a multi-scale framework. *Landsc Ecol* 34:105–117

- Burnham KP, Anderson DR (2002) Model selection and multimodel inference. Springer, New York
- Carle RJ (2006) Factors affecting nest survival of three species of migrant songbirds in the Greater Yellowstone Ecosystem. Masters Thesis, University of Montana
- Cushman S (2007) Research agenda for integrated landscape modeling. US Department of Agriculture, Forest Service, Rocky Mountain Research Station
- Cushman SA, Wasserman TN (2018). In: Humphries G, Magness DR, Huettmann F (eds) Landscape Applications of machine learning: comparing random forests and logistic regression in multi-scale optimized predictive modeling of American Marten occurrence in Northern Idaho. Springer International Publishing, Cham, pp 185–203
- DeVoe JD, Garrott RA, Rotella JJ, Challenger SR, White PJ, O'Reilly M, Butler CJ (2015) Summer range occupancy modeling of non-native mountain goats in the greater Yellowstone area. *Ecosphere* 6:art217
- DNP (2004) Submission for Nomination of the Dong Phrayayen–Khao Yai Forest Complex. Bangkok
- Dzialak MR, Olson CV, Webb SL, Harju SM, Winstead JB (2015) Incorporating within- and between-patch resource selection in identification of critical habitat for brood-rearing greater sage-grouse. *Ecol Process* 4:5
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick RJ, Lehmann A, Li J, Lohmann GL, Loiselle AB, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMM, Townsend Peterson A, Phillips JS, Richardson K, Scachetti-Pereira R, Schapire ER, Soberón J, Williams S, Wisz SM, Zimmermann EN (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57
- European Space Agency (2015) 300 m annual global land cover time series from 1992 to 2015. Climate Change Initiative (CCI), European Space Agency (ESA)
- Evans JS, Cushman SA, Theobald D (2014) An ArcGIS Toolbox for Surface Gradient and Geomorphometric Modeling, version 2.0–0
- Evans JS, Murphy MA, Holden ZA, Cushman SA (2011) Modeling species distribution and change using random forest. In: Drew CA, Wiersma YF, Huettmann F (eds) Predictive species and habitat modeling in landscape ecology: concepts and applications. Springer, New York, NY, pp 139–159
- Fisher TJ, Anholt B, Volpe JP (2011) Body mass explains characteristic scales of habitat selection in terrestrial mammals. *Ecol Evol* 1:517–528
- Gonthier DJ, Ennis KK, Farinas S, Hsieh H-Y, Iverson AL, Batary P, Rudolphi J, Tscharnke T, Cardinale BJ, Perfecto I (2014) Biodiversity conservation in agriculture requires a multi-scale approach. *Proc R Soc B Biol Sci* 281:20141358–20141358
- Goodrich JM, Lynam A, Miquelle DG, Wibisono HT, Kawanishi K, Pattanavibool A, Htun S, Tempa T, Karki J, Jhala Y V, Karanth UK (2015) *Panthera tigris*. In: IUCN Red List Threat. Species 2015 e.T15955A50659951
- Goodwin BJ, Fahrig L (1998) Spatial scaling and animal population dynamics. In: Peterson DL, Parker VT (eds) Ecological scale: theory and applications. Columbia University Press, New York, pp 193–206
- Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2013) High-resolution global maps of 21st-century forest cover change. *Science* 392:850–853
- Harrell FEJ (2018) Harrell Miscellaneous (Hmisc). Comprehensive R Archive Network (CRAN), Version 4.1-1
- Hearn AJ, Cushman SA, Ross J, Goossens B, Hunter LTB, Macdonald DW (2018) Spatio-temporal ecology of sympatric felids on Borneo. Evidence for resource partitioning? *PLoS One* 13:e0200828
- Hebblewhite M, Miquelle DG, Robinson H, Pikunov DG, Dunishenko YM, Aramilev VV, Nikolaev IG, Salkina GP, Seryodkin IV, Gaponov VV, Litvinov MN, Kostyria AV, Fomenko PV, Murzin AA (2014) Including biotic interactions with ungulate prey and humans improves habitat conservation modeling for endangered Amur tigers in the Russian Far East. *Biol Conserv* 178:50–64
- Hegel TM, Cushman SA, Evans J, Huettmann F (2010) Current state of the art for statistical modelling of species distributions. In: Cushman SA, Huettmann F (eds) Spatial complexity, informatics, and wildlife conservation. Springer, Japan, Tokyo, pp 273–311
- Hutchinson GE (1957) Concluding remarks. *Cold Spring Harb Symp Quant Biol* 22:415–427
- Jarvis A, Reuter HI, Nelson V, Guevara E (2008) Hole-filled seamless SRTM data V4. In: Int. Cent. Trop. Agric. <http://srtm.csi.cgiar.org>
- Johnson AR, Wiens JA, Milne BT, Crist TO (1992) Animal movements and population-dynamics in heterogeneous landscapes. *Landsc Ecol* 7:63–75
- Kafley H, Gompper ME, Sharma M, Lamichane BR, Maharjan R (2016) Tigers (*Panthera tigris*) respond to fine spatial-scale habitat factors: occupancy-based habitat association of tigers in Chitwan National Park, Nepal. *Wildl Res* 43:398
- Kanagaraj R, Wiegand T, Kramer-Schadt S, Anwar M, Goyal SP (2011) Assessing habitat suitability for tiger in the fragmented Terai Arc Landscape of India and Nepal. *Ecography* 34:970–981
- Karanth KU, Chundawat RS (2002) Ecology of the tiger: implications for population monitoring. In: Karanth KU, Nichols JD (eds) Monitoring tigers and their prey: a manual for researchers, managers, and conservationists in tropical Asia. Centre for Wildlife Studies, Bangalore, pp 9–22
- Kittle AM, Watson AC, Cushman SA, Macdonald DW (2018) Forest cover and level of protection influence the island-wide distribution of an apex carnivore and umbrella species, the Sri Lankan leopard (*Panthera pardus kotiya*). *Biodivers Conserv* 27:235–263
- Krishnamurthy R, Cushman SA, Sarkar MS, Malviya M, Navreen M, Johnson JA, Sen S (2016) Multi-scale prediction of landscape resistance for tiger dispersal in central India. *Landsc Ecol* 31:1355–1368
- Levin SA (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture *Ecol* 73:1943–1967

- Lynam A, Nowell K (2011) *Panthera tigris* ssp. *corbetti*. In: IUCN Red List Threat. Species 2011 e.T136853A4346984. <https://doi.org/10.2305/IUCN.UK.2011-2.RLTS.T136853A4346984.en>. Accessed 2 Nov 2018
- Macdonald DW, Bothwell HM, Hearn AJ, Cheyne SM, Haidir I, Hunter LTB, Kaszta Z, Linkie M, Macdonald EA, Ross J, Cushman SA (2018) Multi-scale habitat selection modeling identifies threats and conservation opportunities for the Sunda clouded leopard (*Neofelis diardi*). *Biol Conserv* 227:92–103
- Macdonald DW, Bothwell HM, Kaszta Z, Ash E, Bolongon G, Burnham D, Can OE, Campos-Arceiz A, Channa P, Clements GR, Hearn AJ, Hedges L, Htun S, Kamler JF, Kawanishi K, Macdonald EA, Mohamad SW, Moore J, Naing H, Onuma M, Penjor U, Rasphone A, Mark Rayan D, Ross J, Singh P, Tan CKW, Wadey J, Yadav BP, Cushman SA (2019) Multi-scale habitat modelling identifies spatial conservation priorities for mainland clouded leopards (*Neofelis nebulosa*). *Divers Distrib*. <https://doi.org/10.1111/ddi.12967>
- Martin AE, Fahrig L (2012) Measuring and selecting scales of effect for landscape predictors in species–habitat models. *Ecol Appl* 22:2277–2292
- Mashintonio AF, Pimm SL, Harris GM, van Aarde RJ, Russell GJ (2014) Data-driven discovery of the spatial scales of habitat choice by elephants. *PeerJ* 2:e504
- Mateo-Sánchez MC, Balkenhol N, Cushman S, Pérez T, Domínguez A, Saura S (2015) A comparative framework to infer landscape effects on population genetic structure: are habitat suitability models effective in explaining gene flow? *Landsc Ecol* 30:1405–1420
- Mateo-Sánchez MC, Cushman SA, Saura S (2014) Scale dependence in habitat selection: the case of the endangered brown bear (*Ursus arctos*) in the Cantabrian Range (NW Spain). *Int J Geogr Inf Sci* 28:1531–1546
- Mayor S, Schneider D, Schaefer J, Mahoney P (2009) Habitat selection at multiple scales. *Ecoscience* 16:238–247
- McGarigal K (2018) BIOSTATS. Department of Environmental Conservation, University of Massachusetts. v. 9 February 2018
- McGarigal K, Cushman SA (2002) Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. *Ecol Appl* 12:335–345
- McGarigal K, Cushman S, Neel MC, Ene E (2012) FRAGSTATS: spatial pattern analysis program for categorical maps. University of Massachusetts, Amherst
- McGarigal K, Wan HY, Zeller KA, Timm BC, Cushman SA (2016) Multi-scale habitat selection modeling: a review and outlook. *Landsc Ecol* 31:1161–1175
- Nelder JA, Wedderburn RWM (1972) Generalized linear models. *J R Stat Soc Ser A* 135:370–384
- Ngoprasert D, Lynam AJ, Sukmasuang R, Tantipisanuh N, Chutipong W, Steinmetz R, Jenks K, Gale GA, Grassman LI, Kitamura S, Howard J, Cutter P, Cutter P, Leimgruber P, Songsasen N, Reed DH (2012) Occurrence of three felids across a network of protected areas in Thailand: prey, intraguild, and habitat associations. *Biotropica* 44:810–817
- Oksanen J, Minchin PR (2002) Continuum theory revisited: what shape are species responses along ecological gradients? *Ecol Modell* 157:119–129
- Pitman RT, Fattebert J, Williams ST, Williams KS, Hill RA, Hunter LTB, Robinson H, Power J, Swanepoel L, Slotow R (2017) Cats, connectivity and conservation: incorporating data sets and integrating scales for wildlife management. *J Appl Ecol* 54:1687–1698
- R Development Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing. v.3.4.2. Vienna
- Rabinowitz A (1993) Estimating the Indochinese tiger (*Panthera tigris corbetti*) population in Thailand. *Biol Conserv* 65:213–217
- Reddy PA, Cushman SA, Srivastava A, Sarkar MS, Shivaji S (2017) Tiger abundance and gene flow in Central India are driven by disparate combinations of topography and land cover. *Divers Distrib* 23:863–874
- Redfern J V, Ferguson MC, Becker EA, Hyrenbach KD, Good C, Barlow J, Kaschner K, Baumgartner MF, Forney KA, Ballance LT (2006) Techniques for cetacean–habitat modeling. *Mar Ecol Prog Ser* 310:271–295
- Rostro-García S, Kamler JF, Hunter LTB (2015) To kill, stay or flee: the effects of lions and landscape factors on habitat and kill site selection of cheetahs in South Africa. *PLoS One* 10:e0117743
- Rostro-García S, Tharchen L, Abade L, Astaras C, Cushman SA, Macdonald DW (2016) Scale dependence of felid predation risk: identifying predictors of livestock kills by tiger and leopard in Bhutan. *Landsc Ecol* 31:1277–1298
- Royal Forestry Department (2000) Study of the status and database design of natural resources in Khao Yai. Thap Lan, Pang Sida, and Ta Phraya National Parks [Thai]. Royal Forestry Department, Government of Thailand and Geo Asia Co. Ltd., Bangkok
- Shirk AJ, Cushman SA, Waring KM et al (2018) Southwestern white pine (*Pinus strobiformis*) species distribution models project a large range shift and contraction due to regional climatic changes. *For Ecol Manage* 411:176–186
- Shirk AJ, Cushman SA, Waring KM, Wehenkel CA, Leal-Sáenz A, Toney C, Lopez-Sanchez CA (2014) Spatiotemporal variation in resource selection: insights from the American marten (*Martes americana*). *Ecol Appl* 24:1434–1444
- Smith JLD (1993) The role of dispersal in structuring the chitwan tiger population. *Behaviour* 124:165–195
- Sunarto S, Kelly MJ, Parakkasi K, Klenzendorf S, Septayuda E, Kurniawan H (2012) Tigers need cover: multi-scale occupancy study of the big cat in Sumatran forest and plantation landscapes. *PLoS ONE* 7:e30859
- Sunquist M (2010) What is a tiger? Ecology and behavior. In: Tilson R, Nyhus PJ (eds) *Tigers of the world*. Elsevier, New York, pp 19–33
- Sunquist M, Karanth UK, Sunquist F (1999) Ecology, behaviour and resilience of the tiger and its conservation needs. In: Seidensticker J, Jackson P, Christie S (eds) *Riding the tiger: tiger conservation in human-dominated landscapes*. Cambridge University Press, Cambridge, pp 5–18
- Thapa K, Kelly MJ (2017) Prey and tigers on the forgotten trail: high prey occupancy and tiger habitat use reveal the importance of the understudied Churia habitat of Nepal. *Biodivers Conserv* 26:593–616
- Thompson CM, McGarigal K (2002) The influence of research scale on bald eagle habitat selection along the lower Hudson River, New York (USA). *Landsc Ecol* 17:569–586

- Timm BC, McGarigal K, Cushman SA, Ganey JL (2016) Multi-scale Mexican spotted owl (*Strix occidentalis lucida*) nest/roost habitat selection in Arizona and a comparison with single-scale modeling results. *Landsc Ecol* 31:1209–1225
- Toews M (2011) Managing human footprint with respect to its effects on large mammals: implications of spatial scale, divergent responses and ecological thresholds. MSc Thesis. University of British Columbia
- Trisurat Y, Pattanavibool A, Gale GA, Reed DH (2010) Improving the viability of large-mammal populations by using habitat and landscape models to focus conservation planning. *Wildl Res* 37:401–212
- UNESCO (2017) Dong Phrayayen-Khao Yai Forest Complex. In: UNESCO World Herit. Cent. <http://whc.unesco.org/en/list/590>. Accessed 27 Nov 2017
- Wan HY, McGarigal K, Ganey JL, Lauret V, Timm BC, Cushman SA (2017) Meta-replication reveals nonstationarity in multi-scale habitat selection of Mexican Spotted Owl. *Condor* 119:641–658
- Wasserman TN, Cushman SA, Wallin DO, Hayden J, Idaho (2012) Multi scale habitat relationships of *Martes americana* in northern. USDA For Serv - Res Pap RMRS-RP, U.S.A. pp 1–21
- Wibisono HT, Linkie M, Guillera-Aroita G, Smith JA, Sunarto, Pusparini W, Asriadi, Baroto P, Brickle N, Dinata Y, Gemita E, Gunaryadi D, Haidir IA, Herwansyah, Karina I, Kiswayadi D, Kristiantono D, Kurniawan H, Lahoz-Monfort JJ, Leader-Williams N, Maddox T, Martyr DJ, Maryati, Nugroho A, Parakkasi K, Priatna D, Ramadiyanta E, Ramono WS, Reddy G V., Rood EJJ, Saputra DY, Sarimudi A, Salampessy A, Septayuda E, Suhartono T, Sumantri A, Susilo, Tanjung I, Tarmizi, Yulianto K, Yunus M, Zulfahmi (2011) Population status of a cryptic top predator: an island-wide assessment of tigers in sumatran rainforests. *PLoS ONE* 6:e25931
- Wiens J (1976) Population responses to patchy environments. *Annu Rev Ecol Syst* 7:81–120
- Wiens JA (1989) Spatial scaling in ecology. *Funct Ecol* 3:385–397
- Zeller KA, McGarigal K, Beier P, Cushman SA, Vickers TW, Boyce WM (2014) Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: pumas as a case study. *Landsc Ecol* 29:541–557

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.