RESEARCH   Open Access

Multi-scale habitat selection and impacts of climate change on the distribution of four sympatric meso-carnivores using random forest algorithm

Tahir Ali Rather[1,2]* , Sharad Kumar[1,2] and Jamal Ahmad Khan[1]

Abstract

Background: The habitat resources are structured across different spatial scales in the environment, and thus animals perceive and select habitat resources at different spatial scales. Failure to adopt the scale-dependent framework in species habitat relationships may lead to biased inferences. Multi-scale species distribution models (SDMs) can thus improve the predictive ability as compared to single-scale approaches. This study outlines the importance of multi-scale modeling in assessing the species habitat relationships and may provide a methodological framework using a robust algorithm to model and predict habitat suitability maps (HSMs) for similar multi-species and multi-scale studies.

Results: We used a supervised machine learning algorithm, random forest (RF), to assess the habitat relationships of Asiatic wildcat (*Felis lybica ornata*), jungle cat (*Felis chaus*), Indian fox (*Vulpes bengalensis*), and golden-jackal (*Canis aureus*) at ten spatial scales (500–5000 m) in human-dominated landscapes. We calculated out-of-bag (OOB) error rates of each predictor variable across ten scales to select the most influential spatial scale variables. The scale optimization (OOB rates) indicated that model performance was associated with variables at multiple spatial scales. The species occurrence tended to be related strongest to predictor variables at broader scales (5000 m). Multivariate RF models indicated landscape composition to be strong predictors of the Asiatic wildcat, jungle cat, and Indian fox occurrences. At the same time, topographic and climatic variables were the most important predictors determining the golden jackal distribution. Our models predicted range expansion in all four species under future climatic scenarios.

Conclusions: Our results highlight the importance of using multiscale distribution models when predicting the distribution and species habitat relationships. The wide adaptability of meso-carnivores allows them to persist in human-dominated regions and may even thrive in disturbed habitats. These meso-carnivores are among the few species that may benefit from climate change.

Keywords: Multiple-scale, Multi-species, Sympatric carnivores, Species distribution modeling, Bandhavgarh, Climate change

* Correspondence: murtuzatahiri@gmail.com
[1]Department of Wildlife Sciences, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India
[2]The Corbett Foundation, 81-88, Atlanta Building, Nariman Point, Mumbai 400021, India

Introduction

The processes that determine the distribution of the species occur at multiple spatial scales (Wiens 1989; Cunningham and Johnson 2006; Thogmartin and Knutson 2007), e.g., the occurrence of a carnivore species in a habitat patch can depend on the factors that influence the prey densities in their home ranges (third- and fourth-order selection; Johnson 1980). The occurrence can also depend on colonization and dispersal opportunities available in habitat patch (e.g., ecological corridors) (Huck et al. 2010; Zemanova et al. 2017). Thus, the overall occurrence of a species is a function of multiple factors, each most influential at different spatial scales. Multi-scale species distribution models can improve the predictive ability compared to single-scale approaches (Cunningham and Johnson 2006). In multiscale habitat modeling, it is also important to select the predictor variables that significantly influence the species (correct predictors) at the appropriate scales. Failing to do so or including the correct predictors at wrong spatial scales can severely bias the results (Bradter et al. 2013).

One of the most commonly followed approaches of scale selection in multiscale distribution models is to calculate the focal mean of predictor variables within different buffer sizes (scales) around the species occurrence records and by means of appropriate regression technique such as generalized linear models (GLMs) to regress each predictor variables against the response for each scale (Steffan-Dewenter et al. 2002; Holland et al. 2004; Gray et al. 2010). The best scale is then selected by the lowest AIC (Akaike information criterion; Burnham and Anderson 2004). However, studies have shown that a single predictor variable can be equally useful in determining the species occurrence at more than one scale (Bradter et al. 2013), particularly when the ecological predictors are spatially autocorrelated (Legendre 1993). The use of AIC to determine the best scale in such cases can be affected by spatially autocorrelated predictors (Lennon 2000; Hoeting et al. 2006).

Bradter et al. (2013) proposed to evaluate all scales of the predictor variables with moderate levels of statistical support against each other, rather than using the single best scale. Thus, they propose using a more robust machine learning algorithm, random forest (RF; Breiman 2001; Liaw and Wiener 2002), to overcome the uncertainties mentioned above in multi-scale distribution models. RF is robust when the ecological variables are highly correlated (Archer and Kimes 2008; Nicodemus et al. 2010) or the species occurrence points are limited (Strobl et al. 2007). RF also has one key advantage compared to traditional regression modeling when selecting the correct variables. RF not only considers the impact of each individual predictor but also considers the multivariate interactions of variables with each other.

Cushman and Wasserman (2018) compared multiple logistic regression and random forest algorithms to study American martens' multi-scale habitat selection (*Martes americana*). They found that RF outperformed the logistic regression approach. Similar studies report the superior ability of random forest than traditional regression approaches (Cushman et al. 2010; Evans et al. 2011; Drew et al. 2010; Rodriguez-Galiano et al. 2012; Schneider 2012; Cushman et al. 2017). Mi et al. (2017) found that RF demonstrated the best model performance, provided better model fit, and achieved better species range maps than the most powerful and commonly used machine learning algorithms in species distribution models such as TreeNet (boosted regression tree model), CART (classification and regression tree), and Maxent (maximum entropy).

Climate change is an emerging challenge in biodiversity conservation (Bellard et al. 2012) and plays an important role in determining the distribution of the species (Gaston 2003). Studies have suggested that many species will show range contraction and lose a substantial portion of their suitable habitats or become locally extinct due to future climate change (Thomas et al. 2004; Warren et al. 2013). Mammals are particularly vulnerable to climate change (Smith 2013). It is estimated that 19% of the total species locally extirpated in US national parks due to climate change impacts belong to the order Carnivora (Burns et al. 2003). However, it is also hypothesized that generalist species by means of ecological plasticity and broader adaptability may show range expansion and may benefit from future climate change (Thomas 2013).

Medium-sized carnivores (meso-carnivores) have broader niches than top carnivores (Tilley et al. 2013) and are known for their role of scavengers and important seed dispersers (DeVuault et al. 2011; Prugh et al. 2009). In the absence of top predators, meso-carnivores can have a strong influence in shaping the ecological communities (Soulé et al. 1988). Mesopredators may act as top predators in places where the apex predators have been extirpated locally in a process called mesopredator release (Soulé et al. 1988). The mesopredator release may result in the trophic cascade effects (Crooks and Soulé 1999; Berger et al. 2008).

Thus estimating the distributional changes of meso-carnivores under future climatic scenarios is of particular interest. This study assesses the multi-scale habitat association of four meso-carnivore species in and around human-dominated regions. We used scale-optimized predictor variables to predict their current and future distribution under low and high representative concentration pathway scenarios (RCP 2.6 and RCP 8.5). The representative concentration pathways (RCPs) are a set of four new pathways that have been developed for

climatic modeling as a basis for long-term and near-term modeling experiments (van Vuuren et al. 2011).

Materials and methods

Study area

Bandhavgarh Tiger Reserve (BTR) is located between 23° 27′ 00′′ to 23° 59′ 50′′ North latitude and 80° 47′ 75′′ to 81° 15′ 45′′ East longitude in the Umaria district of Madhya Pradesh, in central India (Fig. 1). The reserve's core zone includes the Panpatha Wildlife Sanctuary (PWS) in the north and Bandhavgarh National Park (BNP) in the south, spreading over 716 km². The surrounding buffer zone has an area of 820 km², adding the reserve's total size to 1536 km². The study area represents moist deciduous vegetation dominated by sal (*Shorea robusta*) and sal mixed forests. The overall vegetation of the BTR comprises moist peninsular low-level sal forest, northern dry mixed deciduous forest, dry deciduous scrub, dry grassland, and west Gangetic moist mixed deciduous forest (Champion and Seth 1968).

BTR supports a wide variety of faunal assemblages from small invertebrates to the largest bovid in Asia. There are 35 mammalian species and over 250 species of birds in the reserve. Major large carnivore species include tiger (*Panthera tigris*), leopard (*Panthera pardus*), sloth bear (*Melursus ursinus*), Indian wolf (*Canis lupus*), Asiatic wild dogs (*Cuon alpinus*), and striped hyena (*Hyaena hyaena*). Golden jackal (*Canis aureus*), Indian fox (*Vulpes bengalensis*), jungle cat (*Felis chaus*), Asiatic wildcat (*Felis lybica ornata*), rusty-spotted cat (*Prionailurus rubiginosus*), and fishing cat (*Prionailurus viverrinus*) are the medium-sized carnivores in reserve. The reserve falls within the tropical climatic zone with three distinct seasons; summer, monsoon, and winter. The summer season temperature ranges between 40 and 46 °C, with April and May being generally the hottest summer months. During the winter season, the night temperature may drop abruptly to a minimum of 5 °C. There are more than 150 villages located in the buffer zone of the reserve. The estimated population in and around the reserve is between 40,000 to 50,000 heads. Gond and Baiga are the two major tribal communities living in the study area, and the majority of the population depends on the nearby forests for sustenance. Agriculture and livestock rearing are the primary means of livelihood for the local communities. The collection of non-timber forest produce (NTFP) by local communities is common and widespread throughout the reserve.
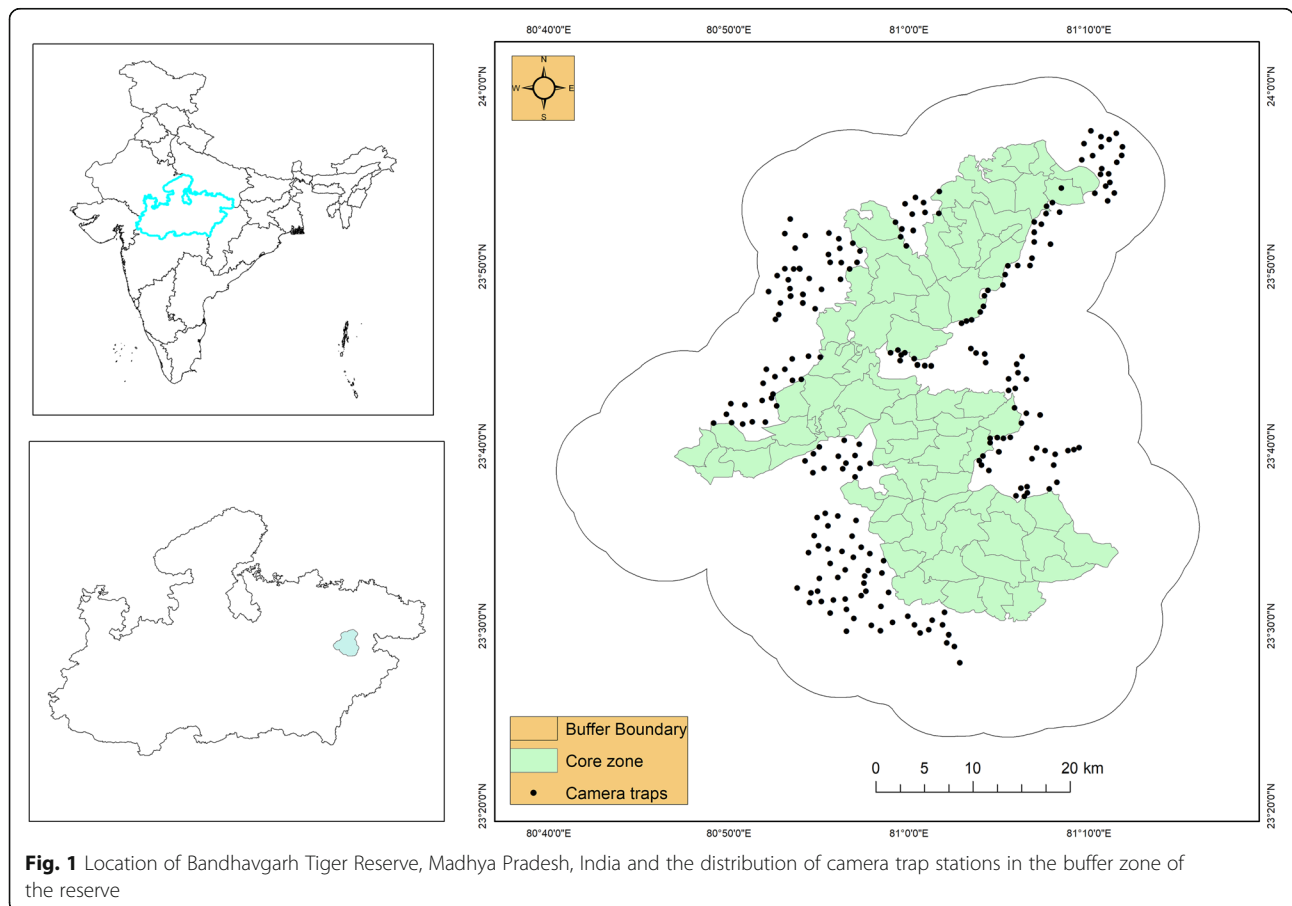
Fig. 1 Location of Bandhavgarh Tiger Reserve, Madhya Pradesh, India and the distribution of camera trap stations in the buffer zone of the reserve

Species occurrence data and spatial autocorrelation

The spatial occurrence data were obtained in the camera trapping survey conducted over an extensive semi-urban buffer zone of the Bandhavgarh Tiger Reserve between January 2016 to May 2016 and September 2016 to December 2016 in the reserve buffer zone. In the second, phase camera traps were placed in territorial forest divisions surrounding the reserve from January 2018 to December 2018. We used the data from 35 pairs of camera traps that yielded 2211 trap nights, and we obtained a total of 544 independent photo-captures of all four species considered in this study. At each camera trap station, we attached a pair of camera traps to the trees along the roads and trails at the average height of 30–40 cm above the ground. We deployed camera traps in 2×2 km grids overlaid in ArcGIS (version 10.3). The spatial data is associated with the inherent bias, commonly referred to as spatial autocorrelation (Dormann et al. 2007). The phenomenon of spatial autocorrelation occurs due to the non-independency of variables sampled at nearby locations from each other (Tobler 1970). Various methods have been developed to account for spatial autocorrelation (Dormann et al. 2007). We used two ways to correct for the spatial autocorrelation in the species occurrence data. First, we implemented spatial filtering using the SDM toolbox (Brown 2014) in ArcGIS (version 10.3) to reduce the spatial bias in the species presence records. The camera trap photo-captures were spatiality rarified at a distance of 1000 m from each other.

Secondly, we tested whether or not the spatially rarified occurrence records assumed the random distribution after implementing spatial filtering by calculating Global Moran's I (Moran 1950) using the Spatial Autocorrelation tool in ArcGIS (version 10.3) (Supplementary S1). After accounting for spatial bias, we retained a total of 55, 60, 49, and 51 spatially rarified occurrence records of Indian fox, golden jackal, jungle cat, and Asiatic wildcat, respectively, for further RF modeling (Supplementary S2, S3, S4, and S5). Lacking the real absence points, we randomly generated pseudo-absence points in ArcGis (version 10.3) in an approximately equal number to the actual occurrence points to deal with the problems arising from unbalanced prevalence (Titeux 2006). The imbalance between the proportion of presence and absence classes causes bias in the model predictions and model fit (Chawla et al. 2003; Chen et al. 2004). In the imbalanced data set, the data's bootstrap is biased toward the majority class (presence or absence), causing the majority class to over-predict. The minority class remains under-predicted. To correct the unbalanced prevalence, we generated twice the random pseudo absence points as the species occurrence records for each species. We then removed the absence points within the buffer

radius of 500 m of the original occurrence points to reduce the number of false negatives (Mateo et al. 2010) and achieve a balanced data set with an approximately equal proportion of presence and absence classes. The buffer distance can be either set arbitrary (as in this study) or based on species attributes (Graham and Hijmans 2006). We used only the occurrence locations collected within the study area for this study. No additional presence records were obtained from any open access data repository platforms such as GBIF (Global Biodiversity Information Facility).

Environmental predictors

This study used 40 environmental predictor variables (Table 1) to predict the species habitat relationships. The predictor variables were grouped into five broad categories: climatic, topographic, landscape composition, vegetation, and human-influenced. The bioclimatic variables were obtained from the WORLDCLIM database (www.worldclim.org). We removed highly correlated predictor variables ($|r| > 0.70$) using r package "rfUtilities" to avoid multi-colinearity among predictor variables (Dormann et al. 2013) (Fig. 2).

We obtained the study area's digital elevation map from the Shuttle Radar Topography Mission (SRTM) elevation database and resampled at 90-m resolution (http://srtm.cs.cgiar.org). Slope, aspect, and topographic ruggedness index was derived from the elevation layer using surface analysis tools in the Spatial Analyst toolbox in ArcGIS (10.3) at the spatial resolution of 90 m. The land use land cover map (LULC) (resampled at 90 m spatial resolution) corresponding to the year 2005 was obtained from the Indian Institute of Remote Sensing (IIRS, http://iirs.gov.in) and reclassified into nine land use categories. The nine categorical habitat variables were derived from the LULC map using the reclassify tool in ArcGIS. During reclassification, the habitat variable of interest was given the value of 1 while keeping all other variables' values at 0. In this way, a reclassified raster layer was derived with the raster values on a continuous scale. Road and river density were calculated using the line density tool in ArcGIS at the spatial scale of 1000, 2000, and 3000 m.

Monthly Normalized Difference Vegetation Index (NDVI) version 6 (MOD13Q1) generated every 16 days available at the spatial resolution of 250 m was obtained from the MODIS website (https://lpdaac.usgs.gov/products/mod13q1v006/) for the year 2013. We reclassified the 23 NDVI layers into three seasons corresponding to summer, wet, and winter seasons and used their average values. The categorical variables were resampled using the nearest neighbor resampling technique, and continuous variables were resampled using bilinear interpolation at the spatial resolution of 90 m in ArcGIS (10.3).

Table 1 The set of 40 predictor variables used in multi-scale habitat modeling of Indian fox, golden jackal, jungle cat, and Asiatic wildcat at most influential spatial scale. All the variables were resampled at the spatial resolution of 90 m. The LULC layer was obtained from Indian Institute of Remote Sensing for the year 2005

Variable type	Variable	Indian fox	Golden jackal	Jungle cat	Asiatic wildcat
Topographic	Elevation	3000	1000	2500	5000
	Slope	5000	1000	2500	4000
	Aspect	3000	3500	5000	3500
	Terrain roughness	2000	4000	1000	5000
	River density	1000, 2000, 3000	1000, 2000, 3000	1000, 2000, 3000	1000, 2000, 3000
Bio-climatic	Bio1	2500	3500	1500	3500
	Bio2	1000	4000	500	5000
	Bio3	5000	1000	3500	4000
	Bio4	5000	1500	3000	4000
	Bio5	2500	4000	3500	1500
	Bio6	500	3500	4500	4000
	Bio7	3000	500	4500	1000
	Bio8	2000	3500	4000	5000
	Bio9	2000	2000	2000	1500
	Bio10	1500	5000	1000	4500
	Bio11	3500	3000	5000	4500
	Bio12	1500	5000	500	2500
	Bio13	3000	1000	1000	2000
	Bio14	4000	3500	500	1500
	Bio15	5000	4000	5000	500
	Bio16	500	1000	500	5000
	Bio17	500	1000	1500	1000
	Bio18	4000	3000	2000	2000
	Bio19	5000	500	3500	500
	Actual evapotranspiration (summer)	2500	1000	3500	3500
	Actual evapotranspiration (wet)	1500	3000	2500	2000
	Actual evapotranspiration (winter)	3000	5000	2000	5000
Landscape composition	Sal dominated	4500	3000	2500	1500
	Sal mix	3500	3000	2000	3500
	Dry deciduous	5000	2500	5000	2500
	Moist deciduous	4000	2000	3500	1000
	Degraded	1500	4500	3500	5000
	Scrub	4000	5000	5000	5000
Vegetation	NDVI (summer)	5000	4500	5000	500
	NDVI (winter)	3500	2500	2500	2000
	NDVI (wet)	3500	3000	3000	2000
Human influenced	Human settlements	2500	3000	4500	4500
	Human population density	5000	2000	1000	3500
	Road density	1000, 2000, 3000	1000, 2000, 3000	1000, 2000, 3000	1000, 2000, 3000
	Farmlands (croplands)	4000	4500	5000	1000

Predictor variables are classified in five groups (topographic, climatic, landscape composition, vegetation and human influenced). Road and river density were calculated at three different spatial scales (1, 2, 3 km). The future projections of bioclimatic variables were downloaded from WorldClim data website (https://worldclim.org/data/index.html). The actual evapotranspiration data was downloaded from MODIS website (https://modis.gsfc.nasa.gov/data/dataprod/mod16.php). The population density data for India was downloaded from WorldPop website (https://www.worldpop.org/).
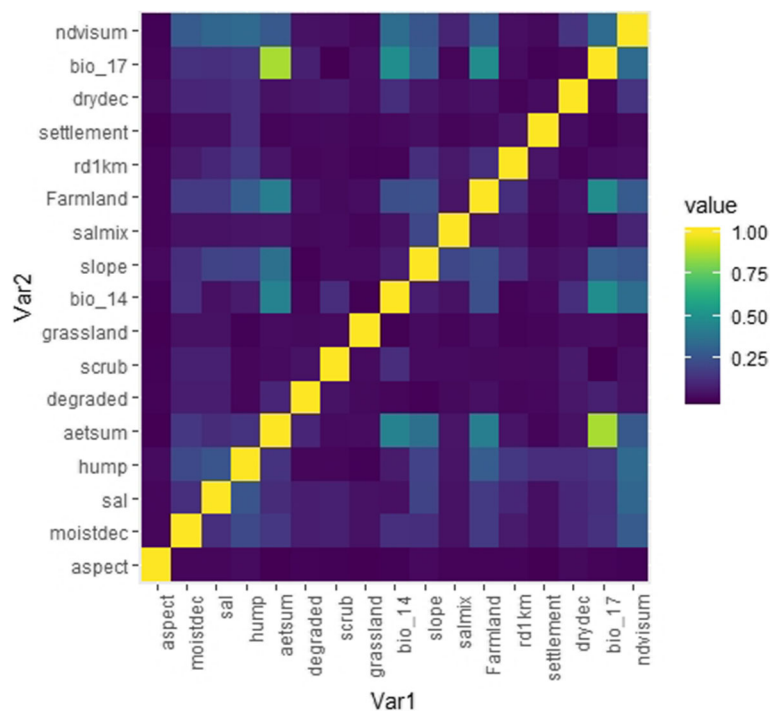
Fig. 2 Multi-colinearity among the predictor variables used in the final multi-scale habitat modeling of all four species

Future climatic data

We modeled the potential distribution of all four species under two representative concentration pathway scenarios (RCP 2.6 and RCP 8.5) developed by Model for Inter-disciplinary Research on Climate change (MIROC5) for the two timelines (the 2050s and 2070s) (Watanabe et al. 2010). These scenarios project the global greenhouse gas emissions based on the assumptions for a wide range of variables such as human population size, global energy consumption, and change in land-use patterns. Global climate models (GCMs) were downloaded from the WordClim website (http://www.worldclim.org/cmip5_30s). The digital layers of GCMs were based on the same bioclimatic layers used to predict the current distribution maps. The climatic models used in this study represent two extreme scenarios of greenhouse gas emissions. The RCP 2.6 assumes that global CO_2 emissions would peak around 2020s and fall to values around zero by 2080. RCP 8.5 is regarded as the worst climatic scenario with higher predicted greenhouse gas emissions. RCP 8.5 assumes that the global CO_2 emissions would increase at a higher rate during the first half of the century and stabilize by 2100. However, the concentrations are three times those in 2000 (Calvente et al. 2009; van Vuuren et al. 2011; Wayne 2013; Rogelj 2012).

Multi-scale data processing

We calculated the focal mean of each predictor variable across ten spatial scales (500–5000 m) surrounding each species occurrence location (presence/pseudo absence) using a moving window analysis with the focal statistic tool in ArcGIS (10.3). Each spatial scale ranging from 500 to 5000 m surrounding each location was used as search radii for calculating the focal mean of all the predictor variables expect road and river density. The road and river density were evaluated at three spatial scales independently (1000, 2000, 3000 m).

Univariate random forest models and scale selection

RF is an ensemble of classification and regression tree (CART) based on the bootstrap aggregation method (also called bagging). The trees are created by drawing samples or sub-samples (bootstrap samples) from the original training data and fitting a single classification or regression tree to each sample. The data that is not part of the bootstrap sample is referred to as out-of-bag (OOB) observations and used to estimate the prediction error rates (OOB error rates). Random forest increases the randomness (diversity) among classification trees by resampling the data with replacement. In this way, RF randomly changes the predictive variable sets over the different tree construction processes. Each classification tree is grown using different bootstrap subsample Xi of the original training data set X (Liaw and Weiner 2002). The bootstrap subsample Xi consists of two-thirds of the observations of the original data set. The observations excluded from the bootstrap subsample are not used in the construction of the ith tree. In the end, each observation of the original training sample is out-of bag

in one-third of the trees constructed. The proportion of the misclassifications over all out-of-bag observations is called an out-of-bag-error rate, usually given in percentage. Following McGarigal et al. (2016) and Cushman et al. (2017), we ran a series of univariate RF models for each predictor variable across ten spatial scales (500–5000 m) to select the most appropriate scale based on the lowest OOB error rates. We used the OOB error rate as a measure of choosing the most suitable predictive spatial scale. In calculating OOB error rates, a training data set was created by sampling with replacement from two-thirds of the data (~ 66%) for each classification tree in a random forest. Each tree was then used to predict the remaining one-third (~ 34%) (out-of-bag sample) of the data, and finally, the OOB error rate was computed as the proportion of times that the predicted class was not the same as the true class (Breiman 2001; Liaw and Wiener 2002). The scale with the minimum OOB error rates was selected as the most influential spatial scale of the predictor variables.

Multi-scale random forest modeling and variable selection

We used scale-optimized predictor variables (scales with lowest OOB error rate) for multi-scale random forest modeling using the package "randomForest" implemented in R (R Core Team 2019). We executed random forest as a regression with the following specifications for each species: 2000 trees (number of bootstrap iterations), ~ 34% data withheld for each tree (out-of-bag [OOB] sample), and m (number of independent metrics permuted at each tree node) optimized to the OOB error estimate following Liaw and Wiener (2002). We followed Murphy et al. (2010) and adopted their methodology to remove redundant metrics, metric selection, evaluating model fit, and test overall model significance. In the first step of multivariate modeling, we used the package "rfUtilities" to account for multi-collinearity among predictor variables by removing the redundant variables using QR matrix decomposition at 0.05 threshold (Becker et al. 1988). The package "rfUtilities" performs QR decomposition on the matrix of predictors to calculate R^2. A predictor is regressed on all other predictors to calculate R^2. If R^2 is less than a particular threshold value, the predictor is said to have failed the test and is subsequently removed from the matrix (see Becker et al. 1988 for details). The multi-collinearity was tested for scale-optimized variables after evaluating the variables for scale selection. Though RF is known to handle a broad set of predictors, however, a large number of predictor variables make ecological interpretation difficult, introduce noise, and decrease the model's explanatory power (Murphy et al. 2010). The RF model output variables are ranked in order of their importance (I) based on

decreasing mean squared error (MSE). However, Murphy et al. (2010) used a new metric model improvement ratio (MIR), which is calculated as [$In/Imax$], where In is the importance of a given metric and $Imax$ is the maximum model improvement score. Our model selection process was based on the procedures briefly described above using MIR. We calculated the MIR scores and then iterated through MIR thresholds (0 to 1 in 0.10 increments) and retained all metrics above this threshold that minimized the model MSE and maximized the percentage of variation explained (Murphy et al. 2010). Removing the redundant variables results in the more parsimonious model with less random noise, and may also improve model OOB error rates (Murphy et al. 2010).

For the variable selection procedures (e.g., Sandri and Zuccolotto 2005; Díaz-Uriarte and de Andrés 2006; Genuer et al. 2010), we followed Genuer et al. (2010) and Murphy et al. (2010) and used permuted variable importance. We identified the most parsimonious model by applying MIR (Murphy et al. 2010). MIR makes use of the permuted variable importance, represented by the mean decrease in model OOB error rate, standardized from zero to one. In this approach, the variables are subset using 0.10 threshold increments, with all variables above this retained for each model (Evans and Cushman 2009; Murphy et al. 2010). This approach is based on the un-scaled permutation importance calculated by permuting each predictor in turn and using the difference in prediction error (OOB error) before and after permutation as a measure of variable importance (Liaw and Wiener 2002; Strobl et al. 2008).

Model validation

We approached model validation using model fit, sensitivity (proportion of observed positives correctly predicted), specificity (proportion of observed negatives correctly predicted), area under the ROC curve, Kappa Statistics, and true skill statistics (TSS). We assessed model fit using the OOB error estimate. The area under the ROC curve (AUC) is a commonly used measure of model performance in the species distribution modeling. Models with AUC values of 0.7–0.9 are considered useful, whereas the values higher than 0.9 are regarded as models with excellent discrimination abilities or high predictive power (Hosmer Jr. et al. 2013; Swets 1988).

Multiscale random forest distribution maps

Following the procedures of univariate random forest models (scale optimization), selection of important variables, and model assessment, we used scale-optimized variables to predict the final distribution maps using the r package "randomForest" in R (Liaw and Weiner 2002). The future distribution maps were predicted using the

same scale-optimized variables. The bioclimatic variables corresponding to the greenhouse gas emission scenarios of RCP 2.6 and RCP 8.5 were used in future prediction maps for the timeline 2050s and 2070s. We converted the continuous distribution probability maps for current and future climatic scenarios into the binary distribution (0,1) using a 0.5 threshold in the final step. The change in the distribution (loss and gain) was calculated in the raster calculator. The distributional change in the output maps was classified as gain (distributional gain in the habitat), loss (distributional loss in the habitat), and stable (no change in the distribution).

Results

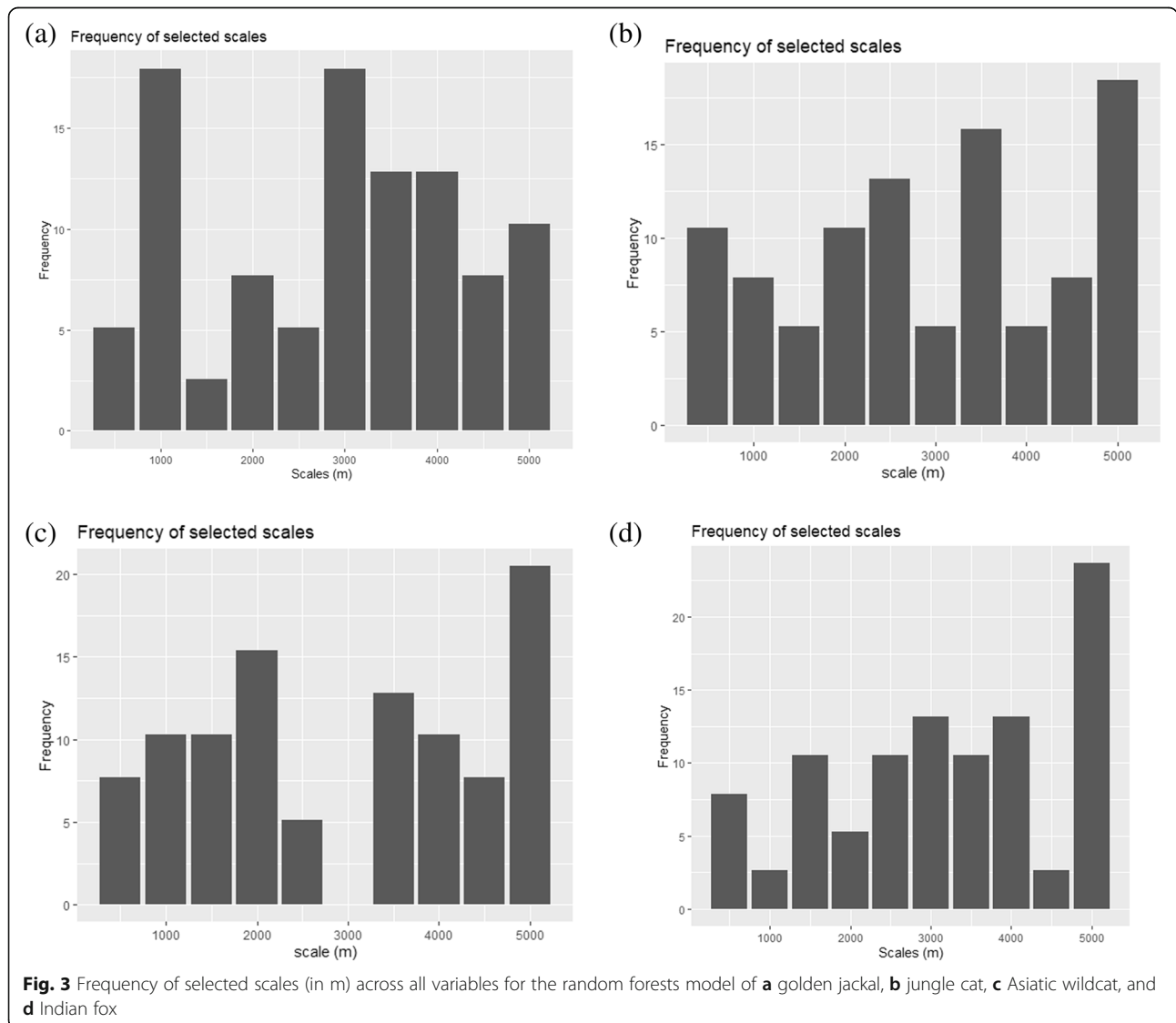### Univariate scaling

A total of ten spatial scales (500–5000 m) for each predictor variable except road and river density were chosen for univariate random forest modeling. Overall, the scales at a broader spatial extent (5000 m) had the highest selection frequency. In jackal, the scales at fine (1000 m) and medium (3000 m) spatial extent occurred more frequently than the broadest scale (Fig. 3a). Jungle cat showed a strong relationship for predictor variables at the broadest scale (Fig. 3b), Asiatic wildcat selected variables more frequently at medium to broader scale (Fig. 3c), and Indian fox perceived habitat variables more regularly at the broadest scale (Fig. 3d).

Multivariate modeling and variable importance

Following the variable selection procedure based on model improvement ratio (MIR) in random forest algorithm, a total of 13, 8, and 12 variables were retained in the final multivariate models of Indian fox, jungle cat,

Fig. 3 Frequency of selected scales (in m) across all variables for the random forests model of **a** golden jackal, **b** jungle cat, **c** Asiatic wildcat, and **d** Indian fox

golden jackal, and Asiatic wildcat (Fig. 3a–d), briefly described in the following components. The selected, scaled variables' contribution is provided in the supplementary information (Supplementary S6).

Indian fox

The aspect was the most important predictor variable, and road density within 1000 m radius was the least important variable based on the variable importance plot (Fig. 4a). The variable importance plot shows variable importance measured as the increased mean square error (MSE), which represents the deterioration of the model's predictive ability when each predictor is replaced in turn by random noise. Higher MSE indicates greater variable importance. Only two predictor variables (degraded forests and actual evapotranspiration in summer) at a small spatial scale (1500–2500 m) were included in the multivariate random forest model of Indian fox (Fig. 4a), two predictor variables (aspect and sal mix forests) were included at medium scales (3000–
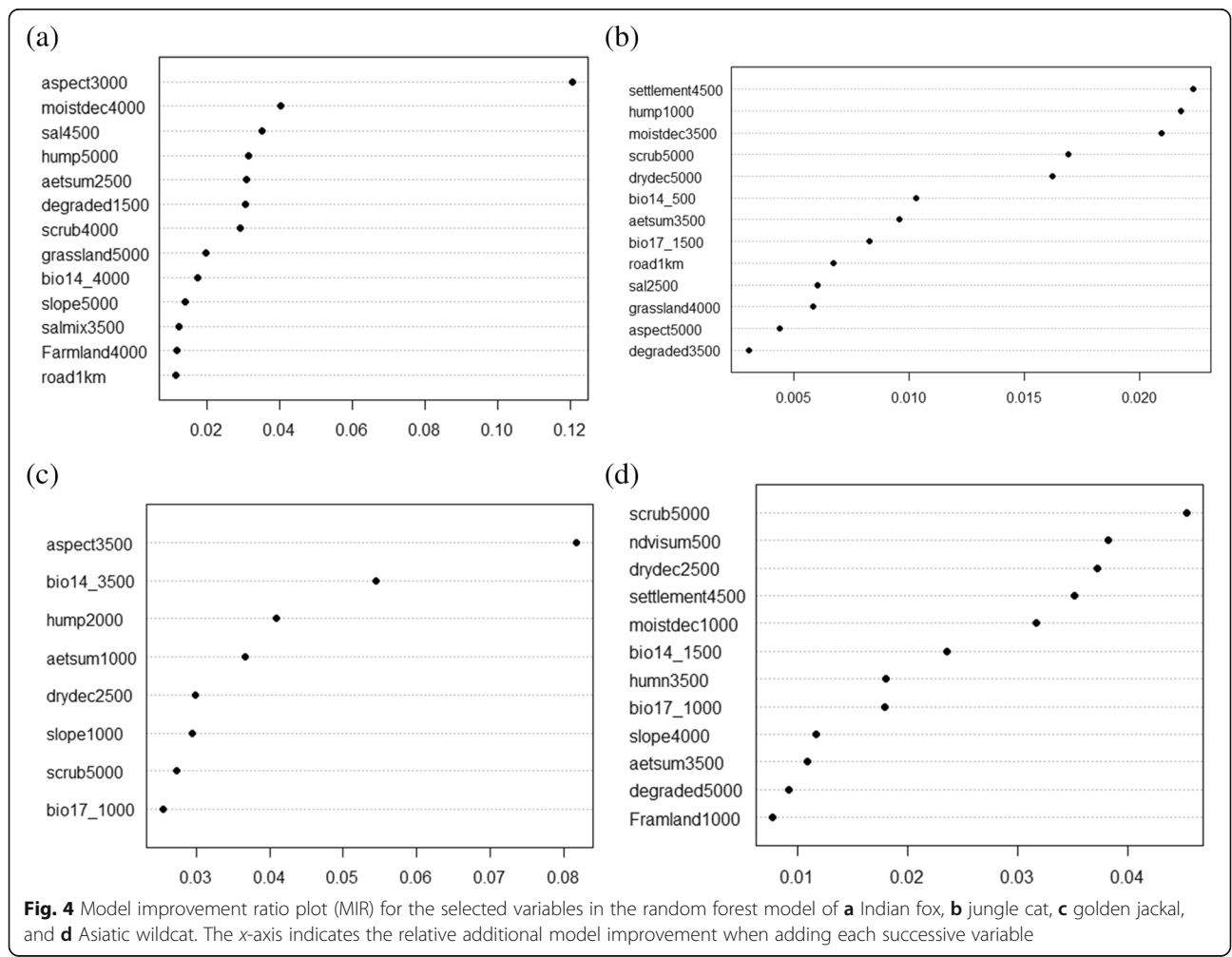
4000 m), and the rest of the eight variables were included at broadest scales (> 4000 m).

Jungle cat

Seven of the thirteen predictor variables were selected at the broadest scale (> 3500 m); three variables, human population density (hump1000), bio14 (bio14_500), and bio17 (bio17_1500), were selected at a small scale (500–1500 m); and sal dominated forest (sal2500) was chosen at medium scale (Fig. 4b). The variable importance plot showed human settlements within the focal radius of 4500 m to be the most important predictor variable of jungle cat occurrence and degraded forests at the spatial scale of 3500 m to be the least important predictor variable (Fig. 4b).

Golden jackal

A total of eight predictor variables were included based on the variable importance plot in the golden jackal's multivariate random forest model (Fig. 4c). Aspect within the focal radius of 3500 m was the most

Fig. 4 Model improvement ratio plot (MIR) for the selected variables in the random forest model of **a** Indian fox, **b** jungle cat, **c** golden jackal, and **d** Asiatic wildcat. The *x*-axis indicates the relative additional model improvement when adding each successive variable

important variable, and bio17 within the focal radius of 1000 m was the least important variable determining the habitat associations of the golden jackal. Four variables in order of their decreasing importance selected at a small spatial scale (1000-2000 m) retained in the final model were human population density (hump2000), actual evapotranspiration in the summer season (aetsum1000), slope (slope1000), and bio17 (bio17_1000). Three variables in order of their decreasing importance at a broader spatial scale (> 3500 m) were aspect (aspect3500), bio14 (bio14_3500), and percentage of scrub habitat patches (scrub5000).

Asiatic wildcat
A total of 13 important variables were selected based on MIR in the final multivariate model of Asiatic wildcat (Fig. 4d). Six variables selected at the broadest scale (> 3500 m) in order of decreasing importance were the percentage of scrub habitats available (scrub5000), human settlement (settlement4500), human population density (humn3500), slope (slope4000), actual evapotranspiration in summer season (aetsum3500), and percentage of degraded forest patches (degraded5000). Five variables selected at small spatial scale (500-2000 m) included Normalized Difference Vegetation Index in summer season

(ndvisum500), percentage of moist deciduous forests (moistdec1000), bio14 (bio14_1500), bio17 (bio17_1000), and percentage of agricultural land (farmland1000).

Partial dependence plots
Partial plots show the marginal effect of single predictor variables included in the respective RF model while keeping the impact of all the other variables on average. The individual species-specific partial dependency plots are provided in Supplementary material S7.
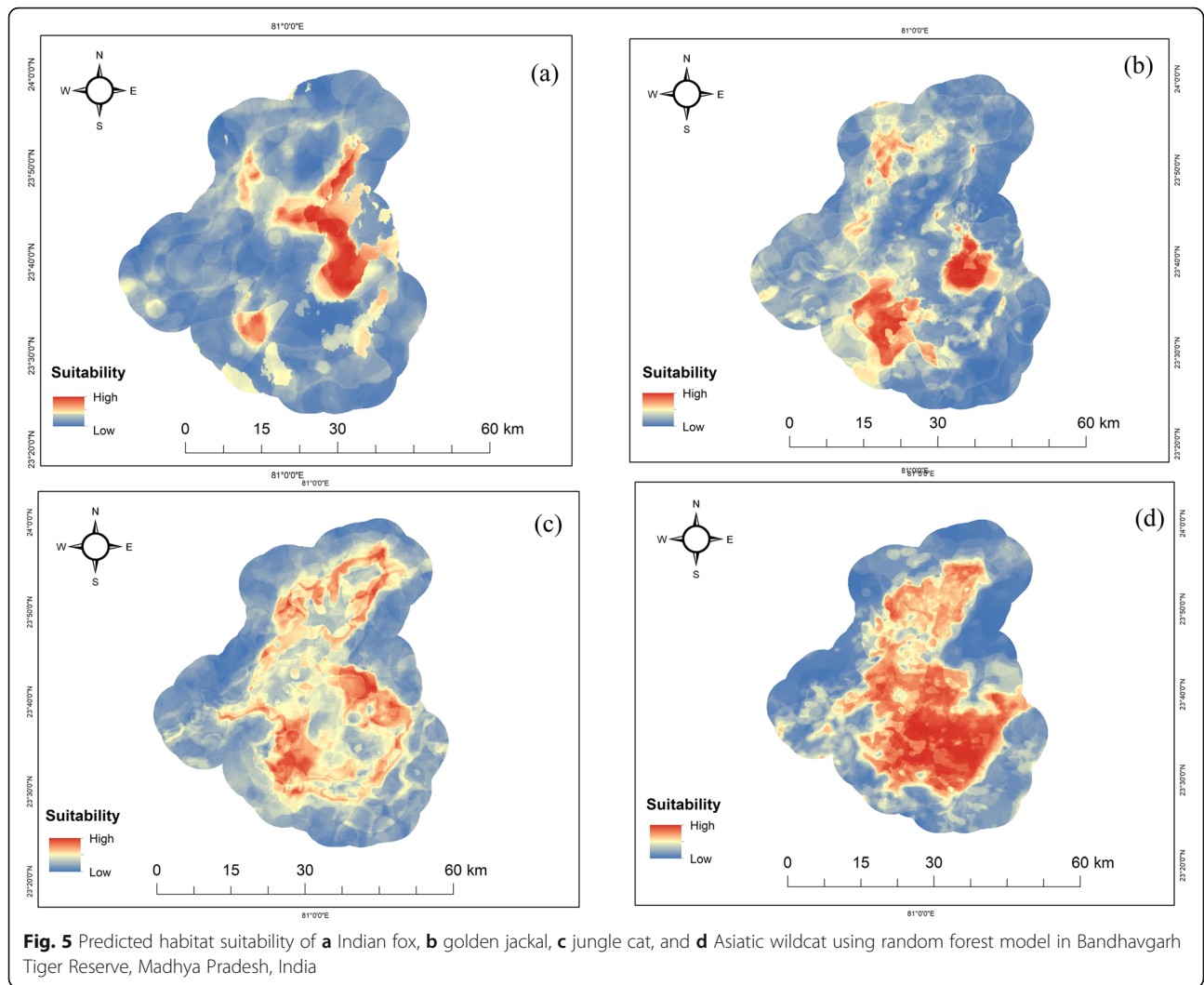
Predicted distribution under current and future climatic scenarios
The distribution of meso-carnivores was predicted using the scale optimized predictor variables selected in the process of variable importance in multiscale random forest models. A total of 39,290.36 (25.57%), 39,211.57 (25.52%), 71,960.76 (46.84%), and 100,015.63 (65.11%) hectares of suitable habitat exists for golden jackal, Indian fox, jungle cat, and Asiatic wildcat, respectively under the current climatic scenario (Table 2).

Suitable areas for Indian fox and jackal were in the eastern and southwestern parts of the reserve, including scrubs, grasslands, degraded forest patches, and a mosaic of croplands and natural vegetation (Fig. 5a, b).

Table 2 Change in the predicted distribution of golden jackal, Indian fox, jungle cat, and Asiatic wildcat in and around Bandhavgarh Tiger Reserve, Madhya Pradesh, India under low (RCP 2.6) and high (RCP 8.5) representative concentration pathway scenarios for the years 2050s and 2070s using the model developed by Model for Inter-disciplinary Research on Climate change (MIROC5)

Species	Scenario	Total stable habitat (ha)	Gain (ha)	Gain (%)	Loss (ha)	Loss (%)	Net gain/Loss (ha)	Net gain/Loss (%)
Jackal	Current	39,290.36						
	2050s RCP 2.6	51,262.55	13,865.52	35.29	1893.33	4.82	11,972.19	30.47
	2050s RCP 8.5	47,478.25	9409.93	23.95	1222.04	3.11	8187.89	− 7.38
	2070s RCP 2.6	41,968.44	5586.23	14.22	2908.15	7.40	2678.08	− 11.60
	2070s RCP 8.5	41,968.44	5586.23	14.22	2908.15	7.40	2678.08	0.00
	Current	39,211.57						
Indian Fox	2050s RCP 2.6	43,530.07	4653.36	11.87	334.86	0.85	4318.50	11.01
	2050s RCP 8.5	42,095.30	3139.79	8.01	256.07	0.65	2883.73	− 3.30
	2070s RCP 2.6	42,149.66	3424.23	8.73	486.14	1.24	2938.09	0.13
	2070s RCP 8.5	38,741.98	551.53	1.41	1021.12	2.60	− 469.59	− 8.08
Jungle cat	Current	71,960.76						
	2050s RCP 2.6	80,318.84	14,355.59	19.95	5997.52	8.33	8358.08	11.61
	2050s RCP 8.5	88,614.67	19,933.16	27.70	3279.25	4.56	16,653.91	10.33
	2070s RCP 2.6	73,828.09	10,725.72	14.90	8858.39	12.31	1867.33	− 16.69
	2070s RCP 8.5	77,264.14	12,665.54	17.60	7362.17	10.23	5303.38	4.65
Asiatic wildcat	Current	10,0015.63						
	2050s RCP 2.6	10,6635.59	7945.21	7.94	1325.25	1.33	6619.96	6.62
	2050s RCP 8.5	10,4189.15	5034.70	5.03	861.18	0.86	4173.52	− 2.29
	2070s RCP 2.6	98,915.71	1448.17	1.45	2548.08	2.55	− 1099.91	− 5.06
	2070s RCP 8.5	10,0382.79	2797.06	2.80	2429.89	2.43	367.16	1.48

Fig. 5 Predicted habitat suitability of **a** Indian fox, **b** golden jackal, **c** jungle cat, and **d** Asiatic wildcat using random forest model in Bandhavgarh Tiger Reserve, Madhya Pradesh, India
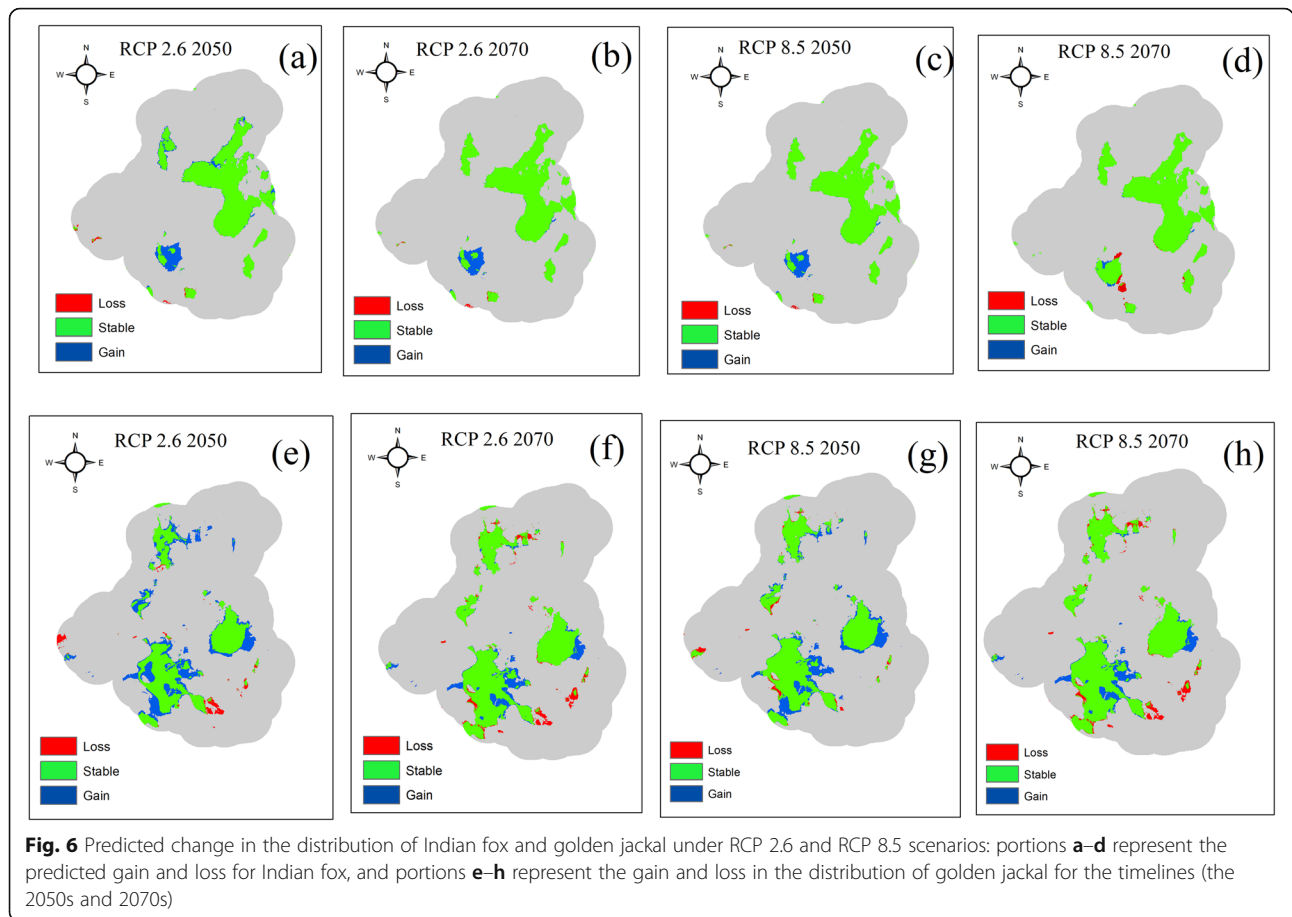
Jungle cat preferred moist and dry deciduous forests, scrubs, grasslands, and degraded habitat patches around human habitations at the reserve's core-buffer interface (Fig. 5c) and most suitable areas for Asiatic wildcat consisted of moist and dry deciduous forests present throughout the reserve (Fig. 5d).

We predicted the change in the distribution of meso-carnivores under the most conservative emission pathway scenario (RCP 2.6) and the worst emission scenario (RCP 8.5). Our model predicted an overall gain in all four species' habitats under all the scenarios considered in this study (Table 3). The highest gain in habitat was predicted for jackal (35%) under RCP 2.6 for the timeline 2050, and the lowest gain was predicted for Indian fox (1.41%) under RCP 8.5 for the timeline 2070 (Table 3). Our model predicted the highest loss in habitat for jungle cat (10.25%) under the RCP 8.5 scenario for the timeline 2070 (Table 3). The gain in the habitat of Indian fox was predicted in the southwest part of the reserve under all emission scenarios (Fig. 6), and the gain in the habitat of the golden jackal was predicted in the

Table 3 Model validation metrics including model OOB error, sensitivity, specificity, Kappa, TSS, AUC, and significance (*p*), for each model

Model	Model OOB error (%)	Sensitivity	Specificity	Kappa	TSS	AUC	*p* value
Asiatic wildcat	0.19	0.733	0.86	0.57	0.61	0.79	0.001
Jungle cat	0.28	0.64	0.79	0.43	0.44	0.71	0.001
Golden jackal	0.16	0.81	0.85	0.66	0.66	0.83	0.001
Indian fox	0.09	0.94	0.82	0.82	0.83	0.91	0.001

Rather et al. Ecological Processes   (2020) 9:60

Page 12 of 17

Fig. 6 Predicted change in the distribution of Indian fox and golden jackal under RCP 2.6 and RCP 8.5 scenarios: portions **a–d** represent the predicted gain and loss for Indian fox, and portions **e–h** represent the gain and loss in the distribution of golden jackal for the timelines (the 2050s and 2070s)

north, east, and southwest of the reserve under RCP 2.6 for the timeline 2050 (Fig. 6).

The gain in the jungle cat's habitat was predicted toward the core areas of the reserve (Fig. 7).

The habitat was predicted to remain stable for golden jackal under RCP 2.6 and RCP 8.5 for the 2070s. Our model predicted the habitat to remain stable for Indian fox under all scenarios, with small amounts of loss predicted under RCP 8.5 for the 2070s. The highest gain in the total stable habitat for jungle cat was predicted under RCP 8.5 for the 2050s, while the Asiatic wildcat habitat was predicted to remain constant under all scenarios considered in this study.

Model assessment

All four models corresponding to Asiatic wildcat, jungle cat, golden jackal, and Indian fox were well supported and significant at $P < 0.001$ (Table 3). The model for the Indian fox was discriminately accurate compared to other species. Models for all four species performed well based on model OOB error rate, though the Indian fox model had the lowest OOB error rate, highest AUC, TSS, and Kappa values (Table 2) compared to the other three species.

Discussion

In this study, we designed our modeling approach to pay explicit attention to selecting the most appropriate spatial scales based on new metrics (OOB error rate) compared to the traditional information criteria such as AIC. We aimed to identify the multi-scale drivers of species distribution using more efficient and accurate modeling techniques. The variables of most importance in the multivariate random forest models of all four species indicate that the landscape composition has a dominant influence on predicting the occurrence of meso-carnivores. We present the first attempt of using multiple scale optimization models to predict the potential change in the distribution of meso-carnivores under the impacts of future climate change using a random forest algorithm. Random forest algorithm is known to be robust to the situations when the ecological variables tend to be highly co-related (Bradter et al. 2013) and outperforms the most commonly used machine learning algorithms in species distribution modeling (Mi et al. 2017). The "ensemble" modeling approach, which combines predictions across different modeling approaches, is generally regarded to perform better than single modeling approaches. However, a recent study compared the

Fig. 7 Predicted change in the distribution of jungle cat and Asiatic wildcat under RCP 2.6 and RCP 8.5 scenarios: portions **a–d** represent the predicted gain and loss for jungle cat, and portions **e–h** represent the gain and loss in the distribution of Asiatic wildcat for the timelines (the 2050s and 2070s)
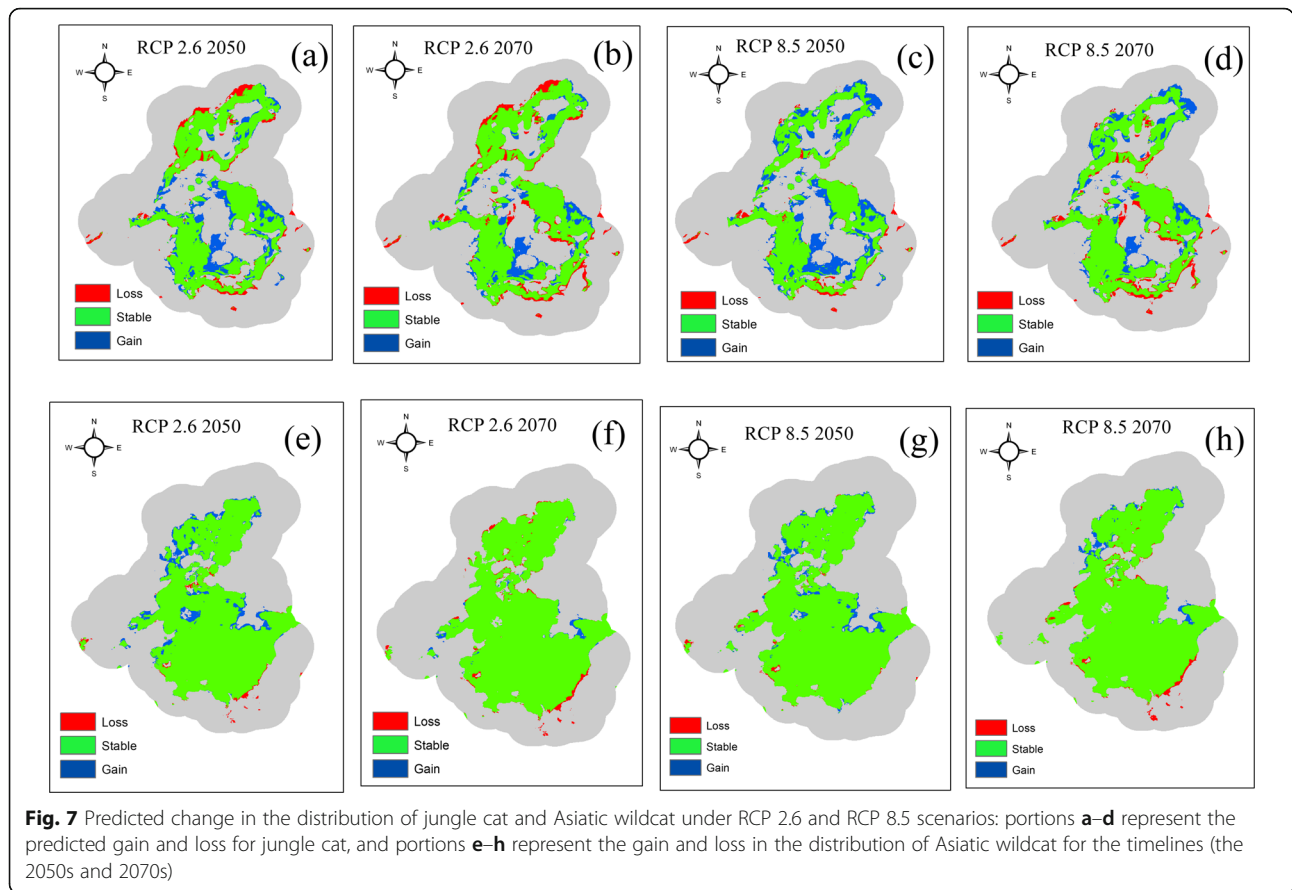
predictive performance of ensemble distribution models to that of individual models and they report no particular benefit to using ensemble modeling approaches over individually tuned models (Hao et al. 2020). The high predictive success and inclusion of scale-optimized variables that drive the distribution of meso-carnivores suggest that these multi-scale niche models may tightly reflect the realized niches of these four species. Thus random forest algorithm may be highly effective at describing the realized niches across complex landscapes. The model improvement ratio (MIR, Murphy et al. 2010) is regarded as an effective method of identifying a more parsimonious set of predictor variables and thus improves the overall model performance (Evans and Cushman 2009).

Scaled variables and partial dependency plots

Our multi-scale and multi-species models selected predictor variables at small (500 m), medium (2000–3000 m), and large spatial scales (> 3500 m). We observed a relatively large difference between the smallest and largest spatial scales selected in the multi-scale models of all four meso-carnivores considered in this study, which indicates the processes determine the distribution of species across multiple spatial scales (Bradter et al.

2013). Our findings agree with the studies that explicitly observe various predictors at specific spatial scales that are most influential in determining the species-habitat associations in multiscale habitat modeling. All four species selected various groups of predictor variables at different spatial scales. The jungle cat responded to the human-influenced variables such as human population density (hump1000), permanent human settlements (settlemnt4500), and degraded habitat patches (degraded3500) at small and broader spatial scales indicating the relatively high tolerance level of jungle cat for the human population within their immediate vicinity. Most of the landscape composition variables were influential at medium and broader spatial scales that may correspond to the third order of habitat selection of jungle cats (Johnson 1980). A higher predicted occurrence of jungle cat was recorded with increasing amounts of moist deciduous forests, sal dominated forests, scrubs, and dry deciduous forests. Rodents are the substantial portion of jungle cats' diet, and habitats such as scrubs, dry deciduous forests, and human settlements may be used for foraging on rodents. The multimodal relationship between jungle cat and degraded forest patches indicates specific tolerance limits at which jungle cat occurrence is recorded highest. The highest predicted

occurrence of the jungle was recorded at high road density within the focal radius of 1 km showing the jungle cat's affinity toward more open areas. Our multi-scale habitat suitability model indicates that dry and moist deciduous forest patches, scrubs, degraded forest patches, forest patches around human settlements, with low precipitation, are the preferred habitats of jungle cats (Kalle et al. 2013).

Asiatic wildcat was previously reported to be distributed exclusively in arid and semi-arid parts of western India. Recently, however, the Asiatic wildcat was recorded in the camera trap survey in the moist deciduous forest of Central India (Rather et al. 2017, 2019), indicating the range expansion in the habitat of Asiatic wildcats in India. Dry and moist deciduous forests were among the most important variables determining its distribution. Among human-influenced variables, human population density, degraded forest patches, and human settlements were most influential in predicting the distribution of wildcat at a larger spatial scale, indicating the tendency of wildcats to perceive the human disturbance at the broadest scale. The foraging sites for Asiatic wildcats may be located in the habitats associated with high rodent abundances such as farmlands, scrubs, and habitat patches near human settlements, which were most influential at small and broad scales.

Indian fox is an opportunistic omnivore generally occurring in dry savannah and open grassland habitats (Vanak and Gompper 2010). Indian fox selected landscape composition variables such as moist deciduous forests (moistdec4000), sal dominated forests (sal4500), scrubs (scrub4000), grasslands (grassland5000), sal mix forests (salmix3500), and farmlands (farmland4000) at broadest scales. In semi-arid landscapes, the grasslands are the important determiners of its distribution (Vanak and Gompper 2010). Our study found habitats such as moist deciduous forests, sal dominated forests, scrubs, sal mix forests, and farmlands were equally important variables of its distribution. Highest predicted occurrences were recorded at higher percentages of degraded forests and farmland.

In contrast, the occurrences of Indian fox declined at higher amounts of grasslands in our study area, conflicting with the findings of Vanak and Gompper (2010). Thus in tropical deciduous forests, the grasslands may not be important variables of fox distribution as in savannah habitats. The highest occurrences of Indian foxes were predicted at high road density values within the focal radius of 1 km, suggesting the affinity for open habitats at the fourth-order of habitat selection.

Golden jackals are generalist and opportunistic omnivores usually associated with open habitats and semi-urban landscapes around human habitations. Our models predicted topographic variables to be significant

predictors determining its distribution in tropical moist deciduous forests. Only two landscape composition variables, dry deciduous forest (drydec2500) within the focal radius of 2500 m and scrubs at the home range level (5000 m), were important variables determining the jackal occurrences. Directionality (aspect), precipitation of the driest month (bio14), actual evapotranspiration in summer (aetsum1000), gentle slopes, and precipitation of driest quarter (bio17_1000) were the most important variables of its distribution.

Predicted distribution and range expansion

All meso-carnivores showed the highest predicted occurrences in habitats represented by significant deforestation, habitat degradation, and human-induced disturbances such as livestock grazing. Our models predicted that range expansion would be related to the climatic and landscape composition predictors under both the low and high emission scenarios. Among the bioclimatic variables, only precipitation-based variables (bio14 and bio17) were significant for the distribution of all four species. All four species showed high predicted occurrences at the low amounts of precipitation of the driest month (bio14) and the driest quarter (bio17), indicating the preference for arid climatic conditions. Under the RCP 8.5 scenario, India's mean warming is predicted to be likely in the range of 1.7–2 °C by the 2030s and 3.3–4.8 °C by the 2080s relative to pre-industrial times. Precipitation under the same scenario is expected to increase by 4% to 5% by 2030s and from 6% to 14% toward the end of the century (Chaturvedi et al. 2012). The range expansion in Indian fox, golden jackal, and wildcat was predicted in degraded forests. In contrast, the range expansion in the jungle was predicted in densely forested habitats toward the reserve's core zone. The high predicted occurrences and the range expansion of meso-carnivores in dry and fragmented habitats are consistent with India's future climatic conditions. Studies investigating the impacts of climate change on forests in India report that about 77% and 68% of India's forest grids are likely to experience a shift in forest types under A2 and B2 scenarios (Ravindranath et al. 2006). The anthropogenic stresses, including livestock grazing, biomass collection for firewood and timber, and the fragmented nature of forests, make the forests more vulnerable to climate change. Thus, it is unlikely that forest area will expand in India (Ravindranath and Sukumar 1998). Meso-carnivores have been reported to expand their range and thrive in human-dominated landscapes (Prugh et al. 2009). Similar results of range expansion in meso-predators are reported from North America, where apex predators have lost 2–76% of their ranges, and meso-carnivores have expanded their range by 60% (Prugh et al. 2009). Mesopredators reportedly broaden their

distribution as the ranges of apex predators decrease (LaPoint et al. 2015). The ranges of tigers are predicted to decrease by 23% in reserve by 2050 (Rather et al. 2020). Habitat fragmentation, deforestation, anthropogenic impacts, and habitat degradation affect large and apex predators negatively compared to mesopredators (Ripple et al. 2014; Šálek et al. 2015). These conditions have been accelerated by recent and ongoing climate changes (Brook et al. 2008). The species can adapt to climate change by either shifting their ranges to remain within the appropriate climatic conditions or by means of adaptations whereby species can alter their behavior or physiology in response to changing temperatures (Foden et al. 2013; Pacifici et al. 2015). The broad adaptability of meso-carnivores to cope with habitat fragmentation and tolerate the human-induced changes enables them to thrive under fragmentation and disturbance conditions. The species with broad niches have benefited from climate change (Pandey and Papeş 2017). Indian fox, golden jackal, jungle cat, and Asiatic wildcat are generalist and opportunistic meso-carnivores with a wide distribution throughout India. They mostly occur in grasslands, scrubs, savannahs, degraded forest patches, croplands, and around human settlements. The proportion of degraded forests, scrubs, and croplands is projected to increase in the future (IPCC 2014), and consequently, we expect the range expansion in meso-carnivores under future climatic scenarios.

Limitations

One of our study's major limitations is that we predicted the changes in the potential distribution of meso-carnivores in a relatively smaller landscape at the size of a single protected area and used limited distributional data. Thus our models may be limited in presenting the complete climatic niches of all the species considered in this study. Though the use of distribution models is frequently used to assess future climate change impacts on potential species distribution, the distribution may also depend on factors such as dispersal abilities, interspecies biotic relationships that were not taken into account. Our study used only one general circulation model (MIROC5) and considered only two emission pathway scenarios, RCP 2.6 and RCP 8.5. The predictions may change using other GCMs. Nevertheless, our study presents the potential response of meso-carnivores to various predictor variables at multiple spatial scales under current and future climatic scenarios.

Conclusion

In this study, explicit attention was given to selecting the most appropriate spatial scales using a robust machine learning algorithm. This study shows that predictor variables were influential at determining the species

occurrence across the range of spatial scales. Thus single-scale approach can severely bias the results and can lead to the wrong management decisions. This study also indicated that human-modified habitats might provide the niches best exploited by habitat generalist species. The expected habitat degradation and expansion in the farmlands to meet the increasing human population's energy requirements under the future climatic scenarios may create the habitats preferred by generalist species. Habitat generalist species are thus likely to expand their range under future climatic scenarios.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13717-020-00265-2.

Additional file 1. Supplementary material S1. The Spatial autocorrelation report as derived from using Spatial Autocorrelation tool in ArcGIS (version 10.3).

Additional file 2. Supplementary material S2. PA_Jackal.

Additional file 3. Supplementary material S3. PA_Jungle.

Additional file 4. Supplementary material S4. PA_Wildcat.

Additional file 5. Supplementary material S5. PA_Indian_Fox.

Additional file 6. Supplementary material S6. Relative importance of variables.

Additional file 7. Supplementary materials S7. Showing the partial dependence plots of Indian fox, golden jackal, jungle cat, and Asiatic wildcat.

Additional file 8. Supplementary materials S8. Predictor variables.

Abbreviations
AIC: Akaike's information criterion; AUC: Area under curve; BTR: Bandhavgarh Tiger Reserve; GCM: General circulation model; GLM: Generalized linear models; HSM: Habitat suitability models; NDVI: Normalized Difference Vegetation Index; RCP: Representative concentration pathway scenario; RF: Random forest; ROC: Resource operating characteristic curve; SDM: Species distribution modeling

Authors' contributions
TAR conceived and designed the study and implemented the analysis. TAR wrote the original manuscript, and SK and JAK reviewed and edited the manuscript. All authors gave final approval for publication. SK and JAK coordinated fieldwork, and SK coordinated field expanses.

Ethics approval and consent to participate
Not applicable.

Consent for publication

Not applicable

Competing interests

We declare no conflict of interest.

References

Archer KJ, Kimes RV (2008) Empirical characterization of random forest variable importance measures. Comput Stat Data An 52:2249–2260

Becker RA, Chambers JM, Wilks AR (1988) The new S language: a programming environment for data analysis and graphics. Wadsworth, Pacific Grove, California, USA

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate change on the future of biodiversity. Ecol Lett 15(4):365–377

Berger KM, Gese EM, Berger J (2008) Indirect effects and traditional trophic cascades: a test involving wolves, coyotes, and pronghorn. Ecology 89: 818–828

Bradter U, Kunin WE, Altringham JD, Thom TJ, Benton TG (2013) Identifying appropriate spatial scales of predictors in species distribution models with the random forest algorithm. Methods Ecol Evol 4:167–174. https://doi.org/10.1111/j.2041-210x.2012.00253.x

Breiman L (2001) Random forests. Mach Learn 45:5–32

Brook BW, Sodhi NS, Bradshaw CJ (2008) Synergies among extinction drivers under global change. Trends Ecol Evol 23:453–460

Brown JL (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5(7):694–700

Burnham KP, Anderson DR (2004) Multimodel inference: Understanding AIC and BIC in model selection. Sociol Methods Res 33:261–304

Burns CE, Johnston KM, Schmitz OJ (2003) Global climate change and mammalian species diversity in US national parks. Proc Natl Acad Sci USA 100(20):11474–11477

Calvente ME, Gil C, Sola AJ, Jiménez-Sánchez ML, Rodríguez-Tamayo ML, Mota JF (2009) Can gypsophytes distinguish different types of gypsum habitats? Acta Botanica Gallica 156(1):63–78

Champion HG, Seth SK (1968) A revised survey of the forest types of India. Government of India Press

Chaturvedi R, Joshi J, Jayaraman M, Bala G, Ravindranath N (2012) Multi-model climate change projections for India under representative concentration pathways. Curr Sci 103(7):791–802

Chawla NV, Lazarevic A, Hall LO, Bowyer KW (2003) SMOTEboost: improving prediction of the minority class in boosting. In: 7th European Conference on Principles and Practice of Knowledge Discovery in Databases, pp 107–119

Chen C, Liaw A, Breiman L (2004) Using random forest to learn imbalanced data. http://oz.berkeley.edu/users/chenchao/666.pdf.

Crooks K, Soulé M (1999) Mesopredator release and avifaunal extinctions in a fragmented system. Nature 400:563–566. https://doi.org/10.1038/23028

Cunningham MA, Johnson DH (2006) Proximate and landscape factors influence grassland bird distributions. Ecol Appl 16:1062–1075

Cushman SA, Gutzwiller K, Evans JS, McGarigal K (2010) The gradient paradigm: a conceptual and analytical framework for landscape ecology. In: Cushman SA, Huettman F (eds) Spatial complexity, informatics, and wildlife conservation. Springer, Tokyo, pp 83–108

Cushman SA, Macdonald EA, Landguth EL, Halhi Y, Macdonald DW (2017) Multiple-scale prediction of forest-loss risk across Borneo. Landscape Ecol 32: 1581–1598

Cushman SA, Wasserman TN (2018) Landscape applications of machine learning: comparing random forests and logistic regression in multi-scale optimized predictive modeling of American marten occurrence in northern Idaho, USA. In: Humphries G, Magness D, Huettmann F (eds) Machine Learning for Ecology and Sustainable Natural Resource Management. Springer, Cham, pp 185–203

Díaz-Uriarte R, de Andrés SA (2006) Gene selection and classification of microarray data using random forest. BMC Bioinformatics 7:3

Dormann CF et al (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628

Dormann CF et al (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46

Drew CA, Wiersma YF, Huettmann F (eds) (2010) Predictive species and habitat modeling in landscape ecology: concepts and applications. Springer Science & Business Media, New York

Evans JS, Cushman SA (2009) Gradient modeling of conifer species using random forests. Landscape Ecol 24:673–683

Evans JS, Murphy MA, Holden ZA, Cushman SA (2011) Modeling species distribution and change using random forest. In: Drew CA (ed) Predictive species and habitat modeling in landscape ecology: concepts and applications. Springer, New York

Foden WB et al (2013) Identifying the world's most climate change vulnerable species: a systematic trait-based assessment of all birds, amphibians and corals. PLoS ONE 8(6):e65427

Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press, London

Genuer R, Poggi JM, Tuleau-Malot C (2010) Variable selection using random forests. Pattern Recogn Lett 31:2225–2236

Graham CH, Hijmans RJ (2006) A comparison of methods for mapping species richness. Glob Ecol Biogeogr 15:578–587

Gray TNE, Phan C, Long B (2010) Modelling species distribution at multiple spatial scales: gibbon habitat preferences in a fragmented landscape. Anim Conserv 13:324–332

Hao T, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2020) Testing whether ensemble modelling is advantageous for maximizing predictive performance of species distribution models. Ecography 43:549–558

Hoeting JA, Davis RA, Merton AA, Thompson SE (2006) model selection for geostatistical models. Ecol Appl 16:87–98

Holland JD, Bert DG, Fahrig L (2004) Determining the spatial scale of species' response to habitat. BioScience 54:227–233

Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression. Wiley, USA. https://doi.org/https://doi.org/10.1002/9781118548387

Huck M et al (2010) Habitat suitability, corridors and dispersal barriers for large carnivores in Poland. Acta Theriol 55:177–192

IPCC (2014) Climate change 2014: synthesis report. In: Pachauri RK, Meyer LA (eds) Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. IPCC, Geneva, Switzerland, p 151

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

Kalle R, Ramesh T, Qureshi Q, Sankar K (2013) Predicting the distribution pattern of small carnivores in response to environmental factors in the Western Ghats. PLoS ONE 8(11): e79295. https://doi.org/https://doi.org/10.1371/journal.pone.0079295

LaPoint S, Belant J, Kays R (2015) Mesopredator release facilitates range expansion in fisher. Anim Conserv 18(1):50–61

Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. Ecography 23:101–113

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18–22

Mateo RG, Croat TB, Felicísimo ÁM, Muñoz J (2010) Profile or group discriminative techniques? Generating reliable pseudo-absences and target-group absences from natural history collections. Divers Distrib 16:84–94

McGarigal K, Wan HY, Zeller KA, Timm BC, Cushman SA (2016) Multi-scale habitat modeling: a review and outlook. Landscape Ecol 31:1161–1175

Mi C, Huettmann F, Guo Y, Han X, Wen L (2017) Why choose random forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. PeerJ 5:e2849. https://doi.org/10.7717/peerj.2849

Moran PAP (1950) Notes on continuous stochastic phenomena. Biometrika 37:17–23

Murphy MA, Evans JS, Storfer A (2010) Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. Ecology 91:252–261

Nicodemus KK, Malley JD, Strobl C, Ziegler A (2010) The behaviour of randomforest permutation-based variable importancemeasures under predictor correlation. BMC Bioinformatics 11:110

Pacifici M et al (2015) Assessing species vulnerability to climate change. Nat Clim Change 5:215–224

Pandey R, Papeş M (2017) Changes in future potential distributions of apex predator and mesopredator mammals in North America. Reg Environ Change 18(4):1223–1233. https://doi.org/10.1007/s10113-017-1265-7

Prugh LR, Stoner CJ, Epps CW, Bean WT, Ripple WJ, Laliberte AS, Brashares JS (2009) The rise of the mesopredator. Bioscience 59(9):779–791

R core Team (2019) R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. URL https://www.R-project.org/.

Rather TA, Kumar S, Kamat A, Gore K (2019) New record of Asiatic wildcat from Central Indian Landscape. Cat News 70:21

Rather TA, Kumar S, Khan JA (2020) Multi-scale habitat modelling and predicting change in the distribution of tiger and leopard using random forest algorithm. Sci Rep 10:11473. https://doi.org/10.1038/s41598-020-68167-z

Rather TA, Kumar S, Tajdar S, Srivastava RK, Khan JA (2017) First photographic record of Asiatic wildcat in Bandhavgarh TR. Cat News 65:35

Ravindranath NH, Joshi NV, Sukumar R, Saxena A (2006) Impact of climate change on forest in India. Curr Sci 90(3):354–361

Ravindranath NH, Sukumar R (1998) Climate change and tropical forests in India. In: Markham A (ed) Potential Impacts of Climate Change on Tropical Forest Ecosystems. Springer, Dordrecht

Ripple WJ et al (2014) Status and ecological effects of the world's largest carnivores. Science 343(6167):1241484

Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens 67:93–104

Rogelj J, Meinshausen M, Knutti R (2012) Global warming under old and new scenarios using IPCC climate sensitivity range estimates. Nat Clim Change 2(4):248–253

Sandri M, Zuccolotto P (2005) Variable selection using random forests. Data Analysis, Classification and the Forward Search. Springer, Berlin, pp 263–270

Šálek M, Drahníková L, Tkadlec E (2015) Changes in home range sizes and population densities of carnivore species along the natural to urban habitat gradient. Mammal Rev 45:1–14

Schneider A (2012) Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sens Environ 124:689–704

Smith AB (2013) The relative influence of temperature, moisture and their interaction on range limits of mammals over the past century. Glob Ecol Biogeogr 22(3):334–343

Soulé ME, Bolger DT, Alberts AC, Wrights J, Sorice M, Hill S (1988) Reconstructed dynamics of rapid extinctions of chaparral requiring birds in urban habitat islands. Conserv Biol 2(1):75–92

Steffan-Dewenter I, Münzenberg U, Bürger C, Thies C, Tscharntke T (2002) Scale-dependent effects of landscape context on three pollinator guilds. Ecology 83:1421–1432

Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8:25

Swets JA (1988) Measuring the accuracy of diagnostic systems. Science 240(4857):1285–1293

Thogmartin WE, Knutson MG (2007) Scaling local species-habitat relations to the larger landscape with a hierarchical spatial count model. Landscape Ecol 22:61–75

Thomas CD (2013) The Anthropocene could raise biological diversity. Nature 502(7469):7

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, Ferreira de Siqueira M, Grainger A, Hannah L, Hughes L, Huntley B, van Jarsveld AS, Midgley GF, Miles L, Orgeta-Huerta MA, Peterson AT, Philips AL, Williams SE (2004) Extinction risk from climate change. Nature 427(6970):145–148. https://doi.org/10.1038/nature02121

Tilley A, López-Angarita J, Turner JR (2013) Diet reconstruction and resource partitioning of a Caribbean marine mesopredator using stable isotope bayesian modelling. PLoS ONE 8(11):e79560

Titeux N (2006) Modelling species distribution when habitat occupancy departs from suitability. Application to birds in a landscape context. PhD thesis. Université Catholique de Louvain, Louvain-la-Neuve

Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. Econ Geogr 46:234–240

van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey Y, Lamarque J-F, Masui T, Meinshause M, Nakicenovic N, Smith SJ, Rose SK (2011) The representative concentration pathways: an overview. Climatic Change 109:5–31

Vanak AT, Gompper ME (2010) Multi-scale resource selection and spatial ecology of the Indian fox in a human-dominated dry grassland ecosystem. J Zool 281:140–148. https://doi.org/10.1111/j.1469-7998.2010.00690.x

Warren R, VanDerWal J, Price J, Welbergen JA, Atkinson I, Ramirez-Villegas J, Osborn TJ, Jarvis A, Shoo LP, Williams SE, Lowe J (2013) Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. Nature Clim Change 3:678–682

Watanabe M et al (2010) Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. J Climate 23(23):6312–6335

Wayne GP (2013) The beginner's guide to representative concentration pathways. Skeptical Science (25)

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3:385–397

Zemanova MA et al (2017) Impact of deforestation on habitat connectivity thresholds for large carnivores in tropical forests. Ecol Process 6:21

Publisher's Note

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