

Identification of Release Habitat of Captive-bred Mammals Demonstrated for Giant Panda in Sichuan Province, China

Zan Zeng^{a, b}, Haoning Wang^c, Shan Gao^{a, b}, Hein van Gils^{a, b}, Yan Zhou^d,
Liya Huang^e, and Xiaolong Wang^{a, b, *, **}

^a College of Wildlife and Protected Area, Northeast Forestry Univ., Ministry of Education, Harbin, 150040 P. R. China

^b Key Laboratory of Wildlife diseases and Biosecurity Management of Heilongjiang Province, Harbin, 150040 P. R. China

^c School of Geography and Tourism, Harbin University, Harbin, 150086 P. R. China

^d Jiangxi Aviation Vocational and Technical College, Nanchang, 330024 P. R. China

^e Changbai Mountain Academy of Sciences, Antu, 133613 P. R. China

*e-mail: yttuhh@yeah.net

**e-mail: wxlhrb@nefu.edu.cn

Received May 10, 2021; revised July 7, 2021; accepted July 10, 2021

Abstract—Two thirds of the giant panda population in the Sichuan Province of China is at risk of extinction due to habitat fragmentation. Connectivity constraints prevent spontaneous re-population of remaining fragments of their historical habitat. On the other hand, the increasing number of captive-bred giant panda makes release into the wild feasible. A comprehensive approach for the identification of potential release habitat and the number of giant panda required for rewilding is demonstrated. The extent of the uninhabited giant panda habitat in 2013 and in 2060 was established by using the MaxEnt species distribution algorithm, published occurrence points ($n = 1014$) and a broad range of landscape variables namely, climatic, terrain, soil, vegetation and human impact, including mines and roads. We used AUC, SD, kappa, CCI, NMI, the odds ratio, TSS, bootstrap replicates and an alternative preprocessing of predictor variables to validate our model. A least cost path (LCP) between habitat fragments was calculated to identify dispersal corridors required for avoidance of inbreeding. Well-connected, uninhabited habitat patches and their number of home ranges were identified. Considering the net reproduction rate of giant panda, we calculated the number of giant panda to be released annually over a 25 or 50 year period, to fully occupy the available home ranges. We identified 6.900 km² well-connected, uninhabited habitat allowing for the annual release of 45–89 or 10–20 captive-bred giant panda for a 25 or 50 year release project. We suggest that our approach may be used for the reintroduction of other large solitary terrestrial mammals.

Keywords: giant panda, rewilding, multicollinearity, maximum entropy, Least Cost Path, dispersal corridor

DOI: 10.1134/S1062359021130082

1. INTRODUCTION

At present, many populations of large mammals, including giant panda, Amur tiger, Amur leopard and brown bear are supported by release and rewilding projects (Goodrich and Miquelle, 2005; Miquelle et al., 2015; Tosi et al., 2015). Such projects need scientific support in setting release targets. In Sichuan, giant panda (*Ailuropoda melanoleuca*)—panda hereafter—feeds on 32 species of bamboo belonging to 7 genera, including staple food such as *Bashania fangiana*, *Yushania brevipaniculata* and *Chimonobambusa szechuanensis* (Yi and Jiang, 2010). The conservation history of the panda in China can be traced back to the 1970s. Then, the large-scale flowering and subsequent die-off of arrow bamboo (*Fargesia nitida*), the main food of panda, resulted in a large number of deaths (Hu, 2004). As a response, trials on breeding captive panda were undertaken dealing with, among others,

the onset of estrus, mating, domestication, artificial insemination and cub nursing (Ding et al., 2019; Hama et al., 2009; Martin-Wintle et al., 2019). Lately, the number of captive-bred panda has been growing at a rate of about 9.3% per year (Xinhuanet, 2019). Moreover, the Giant Panda National Park established in 2017 (Zhao et al., 2019) is available for release. Despite the great effort, panda is still listed as vulnerable in the current Red List of Threatened Species (IUCN, 2019). The vulnerability is partly brought about by habitat fragmentation that hinders gene exchange and accelerates the decrease the genetic diversity within wild populations (Xiong, 2018). Currently, two thirds of the panda in Sichuan province live in small, unconnected habitat patches at a high risk of extinction (Information Office of the State Council, 2015). The level of inbreeding in wild panda is greater than expected for a solitary mammal (Hu et al., 2017).

In order to optimize the number and genetic diversity of panda, two complementary strategies maybe implemented, namely release and rewilding of captive-bred panda into isolated habitat patches (Zhang and Wei, 2006) and establishment of dispersal corridors between such patches (Yin et al., 2010). Rewilding is an accepted rescue strategy for panda. Rewilding panda can preserve genetic diversity and improve the probability of survival in small isolated panda population (Dai et al., 2020). Previous studies have shown that trained and rewilded pandas gradually approximate that of wild pandas on habitat selection (Zhang et al., 2013) and gut microorganism (Jin et al., 2019), which can be taken as an indicator of guarantee of welfare in rewilding of panda. By effective planning and guidance of local government, the conflicts between pandas and locals are rare and endurable. Rewilding is also highly and socially accepted by the local due to its social, economic and ecological benefits (Ma et al., 2016). Recently, pilot cases of successful rewilding of panda have been reported (An, 2017; Wang, 2017; Yu, 2017), suggesting large-scale rewilding has become feasible. The establishment of corridors between two otherwise isolated habitat patches is a common conservation strategy as well and may mitigate inbreeding depression over long time scales (Buglione et al., 2020; Wang et al., 2014).

The panda is mainly found in the Sichuan, Shanxi and Gansu Provinces. In Sichuan, the occupied habitat covers about 25,800 km². The additional potential or uninhabited habitat is estimated at about 9,100 km². Both increased in extent at 11.8 and 6.3% respectively over the past decades (Sichuan Forestry Department, 2015). The number of panda in the wild reached 1864 animals, mostly ($n = 1387$; 74.4%) in Sichuan (Geng, 2015). The annual growth rate was 1.5% (Geng, 2015). The panda density is low (0.07 animals/km²), suggesting potential for population growth (Sichuan Forestry Department, 2015).

The “Captive Giant Panda Release Project” (Swaigood et al., 2011) is expected to start shortly in Sichuan. However, a target density or target number of panda to be released has not been established. Our study aims at contributing to the Release Project by establishing the extent of the uninhabited, well connected panda habitat and the number of captive-bred panda that should be released annually to populate these habitats. The location and extent of potential habitat patches was estimated by a spatial maximum entropy model (MaxEnt), the uninhabited habitat by subtraction of the known occupied habitat. Then, we calculated Least Cost Paths (LCPs) between the habitat patches. The LCPs may function as dispersal corridors to avoid inbreeding. The number of panda to be released annually was assessed by the known home range (Hu et al., 1985) and net reproduction rate (Hou, 2000).

2. MATERIALS AND METHODS

2.1. Research Area

Our research area is the Sichuan Province (4.86×10^5 km²) in southwest China located between 26°03′–

34°19′ N and 92°21′–108°12′ E (Fig. 1). The main landforms (Fig. 1) are the high plateau in the west (4000–4500 m), the central and a southern mountains (>4000 m) and the basin and hills (1000–3000 m) in the east (Shao et al., 2012).

Needle leaved forests dominate on the high plateau and broadleaved deciduous forests in the basin and on the hills. At forested mid-elevations (1500–3000 m) in the central and southern mountains, bamboo is common as understory in forest and forest openings. At the high plateau with its cold-temperate climate fewer bamboo species occur. The basin and hills as well as the lower elevations of the mountains are situated in a humid, subtropical monsoon climate. Above alpine elevations (>3200–3500 m) in the mountains, grass and shrub dominate. The highest elevations are seasonally snow-capped or even glaciated. *Bashania fangiana*, *Yushania brevipaniculata* and *Chimonobambusa szechuanensis*, as the staple food of pandas, are widely and mixed distributed in various habitats of pandas (Gao et al., 2014; Yi and Jiang, 2010).

Sichuan has nearly 87 million inhabitants. The eastern basin and hills have the highest population density. The mountains show a medium density, concentrated in urban settlements at valley bottoms. The western plateau is sparsely populated. The road network in Sichuan includes 7500 km expressways with a large impact on giant panda (Zhou, 2002). Sichuan is covered by forest (40%), grassland (25%), cultivated land (14%), residential land (4%), industry (2%), orchards (2%), water bodies (2%) and unused land (9%) (Wang and Deng, 2006). Opencast mines (2%) are widely distributed in the central mountains coinciding with panda habitat (Yang, 2009).

2.2. Research Data

The panda presence points ($n = 1014$) cover the period from 2011 to 2014 (Sichuan Forestry Department, 2015). We used the four fundamental environmental predictor categories (Table 1) relevant for habitat modeling of terrestrial macro-fauna, namely climate, terrain, vegetation and human impact (van Gils et al., 2014). We used the climate data from the CHELSA database (Karger et al., 2017) of 1979–2013. The Incoming Solar Radiation (ISR) values were calculated per 30 minutes and summed up per growing season. In Sichuan, spring lasts from Feb. 6 to Mar. 31, summer from Apr. 1 to Sep. 5, autumn from Sep. 6 to Oct. 10, and winter from Oct. 11 to Feb. 5 (Li et al., 2007). All spatial data preprocessing and calculations were done with standard operations in ArcGIS 10.2 and projected in UTM-WGS-1984. Where necessary we resampled to 30 arc-seconds.

2.3. Panda Habitat Modeling

A maximum entropy method (MaxEnt) (Phillips et al., 2017) was selected from among Species Distribution Models (SDMs). Potential pitfalls that may affect the accuracy of the model are spatial autocorrelation

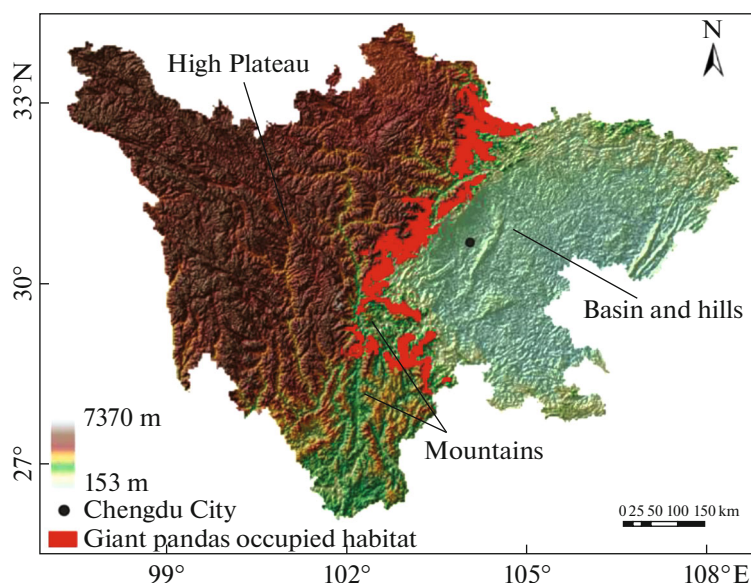


Fig. 1. Location map of Sichuan Province, the occupied habitat of the panda and the main landforms depicted by the digital elevation model (DEM).

of presence points and multicollinearity of predictor variables (Merckx et al., 2011). We minimized spatial autocorrelation by filtering all panda presence points with the SDM Toolbox v1.1c in ArcGIS 10.2. For the first step of the filtering, we used the natural break with a maximum distance of 25 km and a minimum of 5 km (Fekede et al., 2019). In the subsequent spatial rarefying, we set a minimum distance of 13 km between each pair of points. At 13 km, the probability that two points represent the same panda is close to zero (Luo et al., 2017).

Three main methods are used and validated to reduce multicollinearity of predictor variables in spatial modeling, namely, the Principal Component Analysis (PCA) (Farrell et al., 2019; Fekede et al., 2019), the Stepwise Backward Elimination (SBE) (van Gils et al., 2014) and the correlation coefficient elimination (D'Elia et al., 2015). The latter method has been found less optimal (Drake et al., 2006) and has not been considered. The other two have been used in parallel in this paper. Since the number of climate variables is much higher than other variables, we used PCA (SPSS 22.0) to select major climatic predictors (Fekede et al., 2019). We used eigenvalues larger than 1.0 and the scree plot criterion or 'broken stick' stopping rule for PCA in item level factoring (Bernstein, 1988). The climate predictors with factor loading ≥ 0.95 were used for subsequent analysis (Landau and Everitt, 2004). Then, we used PCA for the second time to obtain orthogonal principal component (PC) predictors to reduce multicollinearity (Huang et al., 2018). We used the component score as the coefficient of each standardized predictors to combine the selected climatic predictors and the remaining continuous predictors (human impact, elevation, slope angle, ISR, distance to river). This process uses a

series of orthogonal transformations to re-orient a set of new, uncorrelated variables. The formula for deriving principal components (PCi) is as follows:

$$PC_i = V_{i1}X_1 + V_{i2}X_2 + \dots + V_{ij}X_j + \dots + V_{im}X_m, \quad (1)$$

V_{ij} is the coefficient corresponding to the standardized original variable X_{ij} .

The coefficients' matrix V consists of the eigenvectors from the covariance matrix of the standardized original variables and is obtained by the principal component analysis in SPSS 22.0. The eigenvalue of covariance matrix shows the ability of principal components to synthetically reflect variable X_1, X_2, \dots, X_m . The larger the eigenvalue, the higher the ability of reflecting the original variables. The standardization of variables is achieved by z-score normalization, which is subtraction from the average value and divide the result by the standard deviation (Levi and Rasmussen, 2014).

Finally, we performed Variance Inflation Factor (VIF) analysis on the PCs. A VIF larger than 10 indicates high multicollinearity (Duque-Lazo et al., 2016). The uncorrelated predictor variables and the filtered points served as the data input in MaxEnt. A limitation of PCA is that it is optimized for quantitative data (Vaughan and Ormerod, 2005). Therefore, the categorical variables, soil type and land cover, were fed directly into MaxEnt. The PCs, soil type, land cover and the presence points were used to build and validate the spatial models based on 10 bootstrap replicates. We tested the regularization multiplier (β) between 1 to 5 to determine the optimal value. For the remaining parameters, we kept the default settings. For visualization, the Jenks natural break was used to classify the model output (Liu et al., 2019). Smoothing

Table 1. The environmental predictor variables used in modelling the panda release habitat (Column 1). The source of the variables is provided in column 2. The third column contains the value range of the contiguous variables or for categorical variables the types as found in the research area

Variable	Source	Variable value range or categories (type)
Climate^a	CHELSA	
Monthly <i>P</i>	Ibid.	0 to 545/0 to 581 mm/month
Monthly mean <i>T</i>	Ibid.	−17.5 to 29/−15.5 to 31.3°C
Monthly min <i>T</i>	Ibid.	−25.3 to 25.2/−23.4 to 27.7°C
Monthly max <i>T</i>	Ibid.	−11.9 to 32.6°C(current)/−9.3 to 34.9°C
Bioclimatic (bio1-19)	Ibid.	Supplementary Information Table S.1
Terrain	ASTER-GDEM ^b	
Elevation	Ibid.	−553 to 7845 m a.s.l
Slope angle	Ibid.	0 to 89.5°
ISR-spring	Ibid.	1.2 to 245.0 wh/m ²
ISR-summer	Ibid.	81.4 to 813.9 wh/m ²
ISR-autumn	Ibid.	0.9 to 245.3 wh/m ²
ISR-winter	Ibid.	6.5 to 116.2 wh/m ²
Distance to river	Geospatial data cloud, China ^c	0 to 107.7 km
Human impact		
Population	WorldPop ^d	0.3 to 3940.9 persons/km ²
Distance to road	Geospatial data cloud, China	0 to 63.8 km
Distance to scenic spot	Ibid.	0 to 273.7 km
Distance to settlement	Ibid.	0 to 75.3 km
Categorical		
Soil	SoilGrids ^e	Categorical ^f
Land cover/Vegetation	ESA ^g	Categorical ^h

^a *T* = temperature; *P* = precipitation. Source: <http://chelsa-climate.org/>

^b ISR = Incoming Solar Radiation at ground level. Source: <http://gdem.ersdac.jpacesystems.or.jp/>

^c Source: <http://www.gscloud.cn/>

^d Source: <https://www.worldpop.org/>

^e Source: <https://soilgrids.org/>

^f Soil type: Spodosols, Andisols, Oxisol, Vertisols, Aridisols, Ultisols, Mollisols, Alfisols, Inceptisols and Entisols

^g Source: <https://maps.elie.ucl.ac.be/CCI/viewer/>

^h Land cover: Cropland, Herbaceous, Tree, Shrubland, Grassland, Urban areas, Bare areas, Water bodies and Permanent snow and ice.

for map visualization followed (van Gils et al., 2014). The output was labeled SDM-1.

The process of evaluating or assessing a model is often referred to as “validation” (Minasny et al., 2017). The key component of model training and validation procedures is the criterion which evaluates the model performance. We use threshold dependent and threshold independent criteria. The area under the ROC curve (AUC) is a threshold independent criterion based on plotting the true positives against the false positive fractions for a range of thresholds in prediction probability (Rushton et al. 2004). Currently, the AUC is considered the best criterion for assessing model success for presence/absence data (Austin 2017). AUC is often used as the single criterion for assessment of model performance (Manel and Ormerod 2010; McPherson et al. 2004; Thuiller et al. 2003). As threshold dependent validation measure, we used confusion matrix-based measures including the Kappa test (Liu et al., 2010), correctly classified instances (CCI) (Fielding and Bell, 1997), the nor-

malised mutual information statistic (NMI) (Forbes, 1995), the odds ratio (Fielding and Bell, 1997) and the true skill statistic (TSS) (Allouche et al., 2010). The Kappa statistic and TSS normalise the overall accuracy by the accuracy that might have occurred by chance alone. The percentage of CCI is the rate of correctly classified cells. NMI quantifies the information included in the model predictions compared to that included in the observations. The odds ratio is the ratio of correctly assigned cases to the incorrectly assigned cases. An optimal odds ratio may reach positive infinity, whereas all other criteria are optimal at their maximum, one. Our model has been assessed by all criteria mentioned. Further we compared the average, min, max, median output of the 10 bootstrap replicates with each other. After applying the same visualization process as for the SDM-1, we visually assessed the similarity of the high probability areas of the four outputs. To further validate our model, the coincident area of SDM-1 and the occupied habitat (Sichuan Forestry Department, 2015) was calculated. As addi-

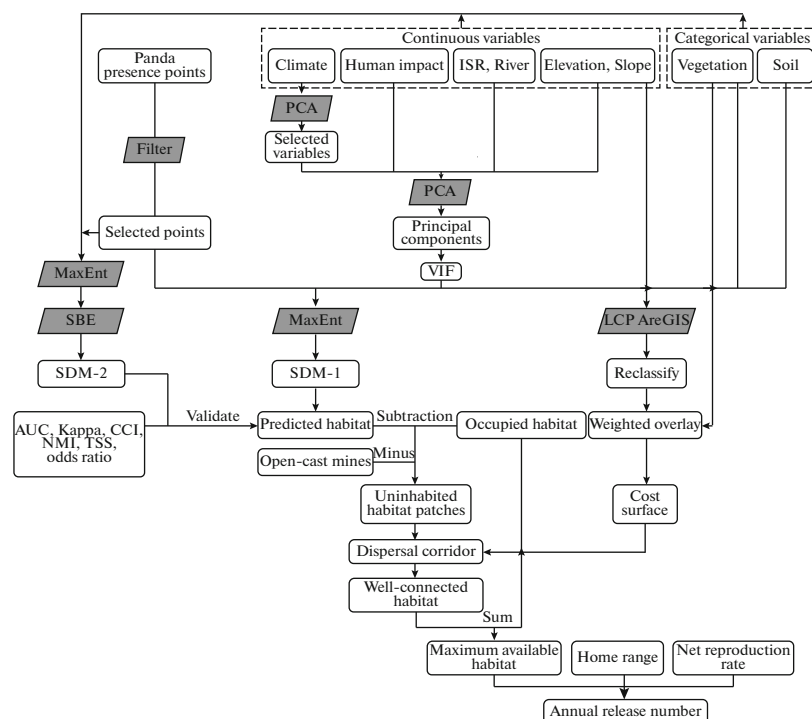


Fig. 2. Flow chart of the procedure used for prediction and assessing the release habitat and the required annual number of released captive-bred panda *SBE* = *Stepwise Backward Elimination*.

tional assessment/validation, we built an alternative model (SDM-2) with the same input data using a posteriori stepwise backward elimination of predictor variables (van Gils et al., 2014) instead of a priori variable selection by PCA (Fig. 2).

Finally, the uninhabited habitat patches were obtained by subtraction of the occupied habitat from the predicted habitat minus the buffered (0.5 km) opencast mines (Sichuan Forestry Department, 2015).

2.4. Dispersal Corridor Analysis

The connectivity of the habitat patches by dispersal corridors was calculated by the LCP approach (Fig. 2). LCP analysis estimates efficient movement routes and costs between pairs of habitat patches based on the suitability of the intervening matrix. LCP analysis is used for many terrestrial mammals, including brown bear (Peters et al., 2015) and cougar (LaRue and Nielsen, 2008). First, we built the cost surface. The cost value of each cell in the grid is based on the mobility of panda in different environments. Vegetation type, elevation and slope-angle raster maps (Table 1: Fig. 2) were used to create the cost surface (Ziółkowska et al., 2012). Elevations between 2000–3000 m, slope-angles from 18–27° and evergreen needle-leaved forest were assigned the lowest cost value (1—lowest resistance, easy to traverse). Raster cells with elevations below 1000 m or above 4000 m, a slope-angle larger than 50°, urban cover and bare cover were assigned the highest

cost value (9—highest resistance, hard to traverse). The intermediate cost values are shown in supplementary information Table S.2. The rating of the three cost factors was added to create a cost surface (LaRue and Nielsen, 2008). LCPs were constructed between each pair of neighbouring habitat patches based on the cost surface (Rayfield et al., 2010). The LCPs were calculated with the Linkage Mapper V 2.0.0 in ArcGIS. The calculated paths were ranked into three classes (high, medium, and low) according to their accumulated costs by means of a K-means algorithm (Hashmi et al., 2017). The uninhabited habitat was divided into potential release regions by size (small/large), location (peripheral to/contiguous with occupied habitat), clustering (scattered/clustered) and connectivity (high/low quality LCP) of habitat patches. Finally, we summated the occupied habitat and the well-connected, uninhabited habitat as maximum available habitat (Fig. 2).

The panda is a solitary animal with a home range of 3.9–6.2 km² showing a negligible overlap (Hu et al., 1985). The number of panda that maybe released was calculated by dividing the extent of the uninhabited, well-connected habitat by the home range. We adopted the average generation period of 12.16 years and the net reproduction rate of 1.0573 (Hou, 2000). The number of panda that could be released annually in the second or fourth generations was calculated with the following formula:

Table 2. The four Principal Components (PC) of the climatic variables, Eigenvalues of each PC and their variance and the cumulative variance of all climatic variables ($n = 67$)

PC	Initial Eigenvalues		
	total	variance %	cumulative %
1	46.1	68.9	68.9
2	14.0	20.8	89.7
3	3.6	5.4	95.1
4	1.4	2.1	97.2

Table 3. The validation criteria of the SDM-1 and the SDM-2

Validation criteria	SDM-1	SDM-2
AUC (area under the ROC curve)	0.96	0.96
SD (standard deviation)	0.003	0.004
Kappa	0.93	0.91
CCI (correctly classified instances)	0.93	0.91
NMI (normalised mutual information statistic)	0.63	0.64
Odds ratio	101.59	112.04
TSS (true skill statistic)	0.82	0.83

$$\text{Annual release number} = \frac{a^n \times 1387 \text{ (Current number of wild panda)}}{\sum_{i=1}^n a^i}. \quad (2)$$

Where: a = net reproduction rate and n = target period (year).

3. RESULTS

Hundred fifty-two (152) panda presence points remained after filtering. The PCA for climatic variables delivered four PCs together accounting for 97.2% of the total variance (Table 2). After PCA, the minimum temperature in April (hereinafter called tmin4, 44.1% contribution), the maximum temperature in June (hereinafter called tmax6, 36.2% contribution) and the temperature seasonality (hereinafter called bio4, 19.7% contribution) were kept for spatial modeling.

Using the component score coefficient (Supplementary Information Table S.3) as a combination of weights, we obtained eight PCs for the three selected climate variables and the other continuous variables. These eight PCs explained more than 90% of the variance and were uncorrelated (VIF 1.0-1.9). The eight PCs plus the categorical variables (soil and land cover) were entered in MaxEnt as environmental layers. With the increase of β , the predicted habitat area is also

increasing. The optimal regularization multiplier for our habitat model is $\beta = 1$. The validation criteria of the SDM-1 and the SDM-2 were nearly identical demonstrating a robust panda habitat model (Table 3).

The overlay pattern of average, min, max, median output of the 10 bootstrap replicates was sufficiently similar (Fig. 3). The standard deviation of the ROC curve (Fig. 4) is small and even. The habitat predicted by the SDM-1 and the SDM-2 is largely the same, further validating our panda habitat model. In addition, our predicted habitat covered 90.5% of the surveyed occupied habitat which may be considered as an independent validation of our model accuracy.

Soil and vegetation (land cover) types contribute together 78.5% to the SDM-1 habitat model and the PCs 21.5%. PC1, PC3 and PC7 are the top contributors among the PCs. The PC1 (8.4%) is related to the ambient temperature and elevation, the PC3 (3.3%) and PC7 (4.2%) represented temperature seasonality and human impact respectively (Supplementary Information Table S.4). The remaining PCs are not considered because of their small contribution. SDM-2 using the SBE method showed that the minimum temperature in June (hereinafter called tmin6, 27% contribution), mean diurnal temperature range (hereinafter called bio2, 25.8% contribution), soil (23.7% contribution) and vegetation (23.5% contribution) were the main predictor variables. Moreover, the probability of panda occurrence is high when the mean diurnal temperature range is small (5.5–6.5°C); the larger the temperature range, the lower the occurrence probability. Further, moderate minimum June temperatures (13–18°C) seem suitable for panda. Spodosols, entisols, broadleaved evergreen forest and, needle-leaved evergreen forests are predictors for panda presence.

The predicted habitat covers a belt from south to north in the middle of the central mountain range in Sichuan (Fig. 5). Habitat patches with medium connectivity occur mainly in the north (Aba), central (Aba), northeast (Bazhong and Dazhou) and the south (Liangshan) of Sichuan, and a small patch in the southeast (Luzhou). The predicted habitat of SDM-1 covers 18.342 km².

The 18.342 km² predicted habitat leaves 18.155 km² after removing the unusable open cast mining areas. The habitat in the Minshan mountains consist of two patches 15 km apart. The predicted habitat the Qionglai and Daxiangling mountains are connected. The predicted habitat patches in the Xiaoxiangling and Liangshan mountains are situated more than 15 km from other patches. The LCP analysis delivered 231 dispersal corridors (Fig. 6). The uninhabited habitat was subdivided into 7 potential release regions (A–G) (Table 4). The well-connected, uninhabited habitat is approximately 6.900 km². Finally, the occupied and well-connected, uninhabited habitat together is about 25.000 km². The occupied and well-connected habitat in Sichuan may ultimately accommodate 4040–6423 individuals. To achieve such numbers, annually 45–89 captive-bred individuals should be released

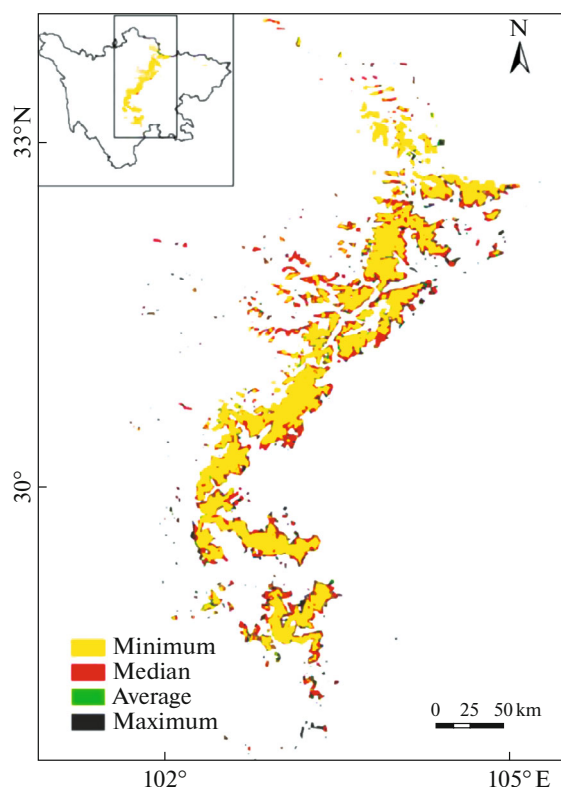


Fig. 3. The average, minimum, maximum and median output of the 10 bootstrap replicates of the SDM-1. In the upper left corner, the box shows the location of the habitat in Sichuan, and the black box shows the enlarged area, which is the main content of this figure.

over 25 years. Alternatively, full rewilding would take 50 years at an annual release of 10–20 captive-bred panda.

4. DISCUSSION

Our finding on the β parameter of MaxEnt is consistent with the characteristics of overfitting (Merckx et al., 2011). This implies that the default setting ($\beta = 1$) of MaxEnt is also correct in our research setting.

The two methods used in this study to reduce multicollinearity have their own advantages and disadvantages. The contribution of the environmental predictor variables were traceable with the stepwise backward elimination (SBE) method. However, the SBE may ignore a unique contribution of an omitted variable that could result in a substantial loss of explanatory power (Carnes and Slade, 1988; James and McCulloch, 1990). However, more recent research did not find such loss, neither within a contiguous research area (Gils et al. 2014) nor in testing the transferability of the model to a disjunctive area (Duque-Lazo et al. 2016). As for the PCA method, the first PC (PC1) contained the most of the data variability of the original variable. The succeeding component (PC2) had the highest variance possible under the constraint that it is orthogonal to the preceding component (Huang et al., 2018). The 8 PCs used in this study can explain more than 90% of the original variables, but it cannot determine the response of the original variables in the model. In this study, SDM-1 and SDM-2 are mutually verified, and the similarity of the results shows that both methods are effective. How these two methods should be selected in a different context requires further research.

The dominant contribution of the two categorical predictor variables vegetation and soil (78.5%) in the SDM-1 over and above climate and elevation comes not as a surprise. A similar finding was reported and explained in a bear distribution study in a mountainous environment (Gils et al. 2014). Therefore, spatial distribution models exclusively based on extrapolated

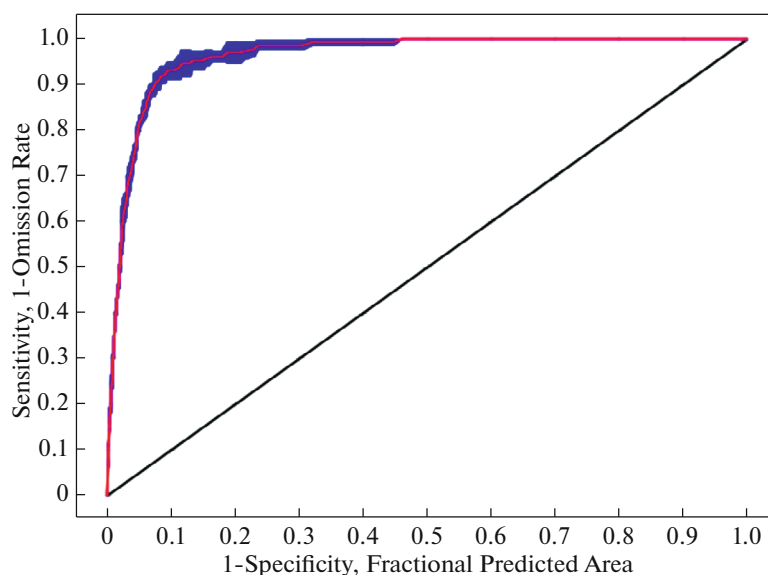


Fig. 4. Receiver operating characteristic (ROC) curve of SDM-1 in red. The blue area is the SD, and the black line the random prediction.

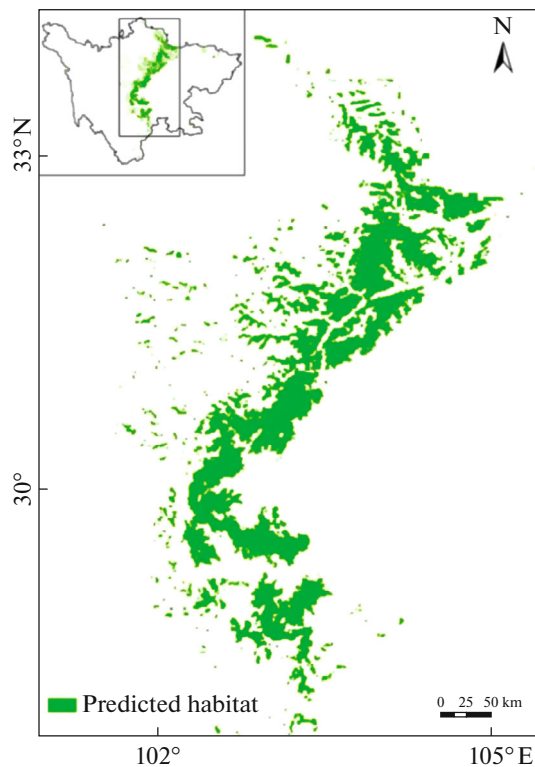


Fig. 5. Predicted panda habitat by the SDM-1 in the Sichuan Province. In the upper left corner, the box shows the location of the habitat in Sichuan, and the black box shows the enlarged area, which is the main content of this figure.

climatic variables (WorldClim) should be interpreted carefully. More so as the current panda study used climatic data (CHELSA) extrapolated explicitly for the particulars of mountainous environments.

The human impact variables (roads, population density, scenic spot, opencast mines) used in this study contributed very little to the predicted habitat. This may be due to the fact that the remaining wild panda live in extensive forested areas without human settlement or isolated farms. Therefore, we did not include three human impact variables in the LCP calculation. However, opencast mines are known to have a substantial, negative local impact (Qi, 2009), as also suggested by the absence of panda records within a radius of 0.5 km of the mines (Sichuan Forestry Department, 2015).

The low contribution of the human population density variable on the model output may result from using global databases on population density and urban land cover in the absence of accessible, comprehensive provincial data. Such global databases imply out of necessity broad generalization. We suggest carrying out a provincial remote sensing study, including rural settlements, rural roads and agrarian land cover/use changes and rerun the habitat model with the more detailed data.

Our release target for pandas is based on the assumption that there is no significant change in habitat over the next 25–50 year. As the future climate change in Sichuan is predicted mainly concentrated at

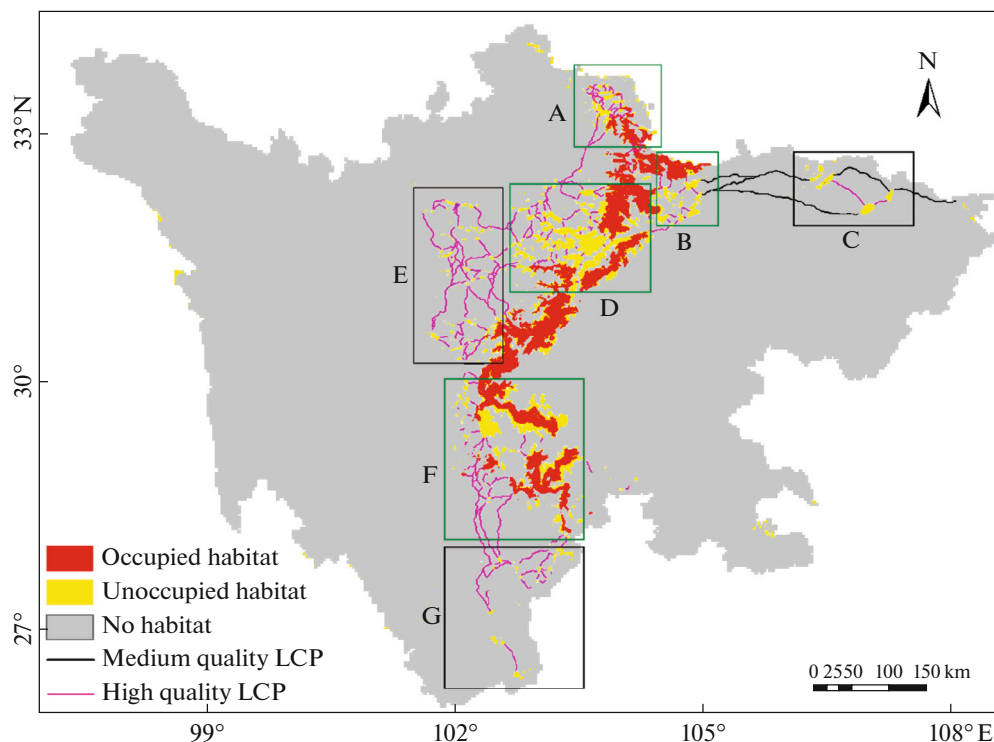


Fig. 6. The current occupied and the uninhabited panda habitat in the Sichuan Province with their medium and high-connectivity LCP dispersal corridors. Black box (A, B, D and F))) = peripheral, small release habitat patches. Green box (C, E and G))) = central large release habitat patches.

Table 4. The predicted uninhabited habitat patches grouped into A-G regions from north to south (Fig. 6) with their pertinent release features and size in km²

Region	Uninhabited habitat patches		Release habitat	Recommendation
	Description	km ²	km ²	
A	<ul style="list-style-type: none"> • Relatively concentrated; 3–13 km apart • High quality LCPs 	26	26	Restore continuous habitat
B	Contiguous with occupied habitat	602	602	Restore continuous habitat
C	Low/medium quality LCPs	243	0	Ignore
D	<ul style="list-style-type: none"> • large; >2000 km² • clustered • contiguous with occupied habitat • high quality LCPs • fragmented by roads/towns 	2103	2103	Restore isolated habitat
E	Isolated by high (>3500) mountains	510	0	Ignore
F	<ul style="list-style-type: none"> • mostly large; >4000 km² • contiguous with occupied habitat • natural dispersal feasible 	4167	4167	Natural dispersal expected
G	<ul style="list-style-type: none"> • scattered; 6–70 km apart • high quality LCPs 	179	0	Ignore
A–G	—	7830	6898	—

the highest elevations region (3800–4500 m) which is above the panda habitat (Lu et al., 2015), we exclude the influence of climate change on the habitat over the next 25–50 year.

Bamboo is the staple food of panda, and it's sufficient supply is the premise that the giant pandas can fully occupy the maximum available habitat in the future. Some studies presented a pessimistic view on the impact of climate change on the extent and species of bamboo (Li et al., 2015). However, but the study used bamboo data from 2003 and only considered climate variables. The most recent data showed an increase of 61.49% of the bamboo covered area in Sichuan Province since 2003 (Information Office of the State Council, 2015). This substantial increase in bamboo over the past decade or so may be due to efforts to restore the panda habitat, including planting bamboo. Therefore, we speculate that bamboo resources can support the expansion and reintroduction of panda habitat.

The potential release targets as provide by our study will neither be limited by birth rate and high mortality of the captive-bred panda population (Li et al., 2017), nor by the number of young captive-bred pandas (Zhao, Zhang et al. 2017). Most female panda will reach the reproductive age (5–25 years) over the next ten years. Moreover, most captive-bred panda give birth to twins every second year (Hou, 2000). The limitation to reach the release target is likely to be the capability of pre-release training of panda.

5. CONCLUSIONS

Our study simulated the habitat of panda in Sichuan Province, and found that there were 6898 km² of habitat patches with good connectivity in addition to the occupied habitat. To fully exploit the available well-connected habitat, 45–89/10–20 captive-bred individuals should be released every year for 25/50 years. Through this example, we construct a two-step PCA method to restrict multicollinearity among a large number of predictor variables. Overall, we have constructed a refined method to predict mammal habitats and release targets, and demonstrated its feasibility with multiple validation. The methods (MaxEnt; LCP), variables (vegetation/land cover, soil, climate, human impact) and parameters (home range; reproductive rate) used in our approach are applicable to many solitary terrestrial mammals and may provide insights helpful in setting feasible release targets.

FUNDING

This work was supported by the National Key Research and Development Program of China (Grant No. 2017YFD0501702); State Key Laboratory of Veterinary Biotechnology Foundation (Grant No. SKLVB201904) and Fundamental Research Funds for the Central Universities (Grant No. 2572017PZ11).

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

AUTHOR CONTRIBUTIONS

Xiaolong Wang conceived and supervised the entire study. Zan Zeng carried out the spatial modeling, the Least Costs Path analysis, interpretation of results and map design. Haoning Wang, Shan Gao, Yan Zhou and Liya Huang contributed to the data collection. Hein van Gils contributed to the spatial modeling and the discussion of methods and results. All authors significantly contributed to the final manuscript.

REFERENCES

- Allouche, O., Tsoar, A., and Kadmon, R., Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS), *J. Appl. Ecol.*, 2010, vol. 43, pp. 1223–1232.
- An, Y. Wilderness release panda “Zhang Xiang” appeared in the nature reserve. <http://www.chinanews.com/sh/2017/04-10/8195713.shtml>. Accessed December 5, 2019.
- Bernstein, I.H., Garbin, C.P., and Teng, G.K., *Applied Multivariate Analysis*, New York: Springer, 1988.
- Buglione, M., Troisi, S.R., Petrelli, S., van Vugt, M., Notomista, T., Troiano, C., et al., The first report on the ecology and distribution of the wolf population in Cilento, Vallo di Diano and Alburni National Park, *Biol. Bull.*, 2020, vol. 47, pp. 640–654.
- Carnes, B. and Slade, N., The use of regression for detecting competition with multicollinear data, *Ecology*, 1988, vol. 69, p. 1266.
- D’Elia, J., Haig, S. M., Johnson, M., Marcot, B.G., and Young, R., Activity-specific ecological niche models for planning reintroductions of California condors (*Gymnogyps californianus*), *Biol. Conserv.*, 2015, vol. 184, pp. 90–99.
- Dai, Q., Li, J., Yang, Y., Li, M., Zhang, K., He, L., et al., Genetic diversity and prediction analysis of small isolated giant panda populations after release of individuals, *Evol. Bioinform.*, 2020, vol. 16, pp. 1–9.
- Ding, Q., Yu, J., Xie, C., Xu, C., Sun, Q., Bi, J., et al., A preliminary study on the maternal behavior of female giant pandas in Shanghai Wild Animal Park, *Acta Theriol. Sin.*, 2019, vol. 39, pp. 302–310.
- Drake, J.M., Randin, C., Guisan, A., Modelling ecological niches with support vector machines, *J Appl Ecol*, 2006, vol. 43, pp. 424–432.
- Duque-Lazo, J., van Gils, H., Groen, T.A., and Navarro-Cerrillo, R.M., Transferability of species distribution models: The case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia, *Ecol Model*, 2016, vol. 320, pp. 62–70.
- Farrell, A., Wang, G., Rush, S., Martin, J., Belant, J., Butler, A., et al., Machine learning of large-scale spatial distributions of wild turkeys with high-dimensional environmental data, *Ecology and Evolution*, 2019, vol. 9, pp. 5938–5949.
- Fekede, R.J., van Gils, H., Huang, L., and Wang, X., High probability areas for ASF infection in China along the Russian and Korean borders, *Transboundary Emerging Dis.*, 2019, vol. 66, pp. 852–864.
- Fielding, A. and Bell, J., A review of methods for the assessment of prediction errors in conservation presence/absence models, *Environ. Conserv.*, 1997, vol. 24, pp. 38–49.
- Forbes, A.D., Classification-algorithm evaluation: five performance measures based on confusion matrices, *J. Clin. Monit.*, 1995, vol. 11, pp. 189–206.
- Gao, Q., Luo, C., Dai, B., Liu, L., Dan, M.A., Physical properties and the impact factors on soil under the bamboo forest in the South Sichuan Bamboo Forests area, *China For. Sci. Technol.*, 2014, vol. 28, pp. 52–55.
- Geng, G., 1864 Wild Giant Pandas Protection in China makes new achievements, *Green China*, 2015, vol. 2, pp. 10–12.
- Goodrich, J.M. and Miquelle, D.G., Translocation of problem Amur tigers *Panthera tigris altaica* to alleviate tiger-human conflicts, *Oryx*, 2005, vol. 39, pp. 454–457.
- Hama, N., Huang, Y.A.N., Kanemitsu, H., Oyama, Y., Ma, Q., Luo, B.O., et al., Artificial insemination and still-birth in a giant panda (*Ailuropoda melanoleuca*), *Jpn. J. Zoo Wildlife Med.*, 2009, vol. 14, pp. 119–123.
- Hashmi, M.M., Frate, L., Nizami, S.M., and Carranza, M.L., Assessing transhumance corridors on high mountain environments by least cost path analysis: the case of yak herds in Gilgit-Baltistan, Pakistan, *Environ. Monit. Assess.*, 2017, vol. 189, p. 488.
- Hou, W., Amendment on the life table and relevant datum indicators of giant panda, *Zool. Res.*, 2000, vol. 21, pp. 361–366.
- Hu, J., Population and conservation of giant pandas in Wolong and Caopo Nature Reserves, *Acta Theriol. Sin.*, 2004, vol. 24, pp. 48–52.
- Hu, J., Schaller, G., and Pan, W., *Wolong’s Giant Panda*, Chengdu: Sichuan Publishing House of Science and Technology, 1985.
- Hu, Y., Nie, Y., Wei, W., Ma, T., Van Horn, R., Zheng, X., et al., Inbreeding and inbreeding avoidance in wild giant pandas, *Mol. Ecol.*, 2017, vol. 26, pp. 5793–5806.
- Huang, C., Wei, H., Rau, J., and Jhan, J., Use of principal components of UAV-acquired narrow-band multispectral imagery to map the diverse low stature vegetation fAPAR, *GIScience Remote Sensing*, 2018, vol. 56, pp. 605–623.
- Information Office of the State Council. State Forestry administration holds the fourth national press conference on the results of giant panda survey. <http://www.scio.gov.cn/xwfbh/gbwxfbh/fbh/Document/1395514/1395514.htm>. Accessed December 5, 2019.
- IUCN, The IUCN Red List of Threatened Species, version 2019-2. <http://www.iucnredlist.org>. Accessed December 5, 2019.
- James, F.C. and McCulloch, C.E., Multivariate analysis in ecology and systematics: panacea or Pandora’s box?, *Annu. Rev. Ecol. Syst.*, 1990, vol. 21, pp. 129–166.
- Jin, L., Zhou, Y., Li, C., Wu, D., Li, G., He, Y., et al., Composition and variation of gut microbiome of trained, preparatory reintroduced, reintroduced and wild giant pandas, *Chin. J. Appl. Environ. Biol.*, 2019, vol. 25, pp. 344–350.
- Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., et al., Climatologies at high resolution for the earth’s land surface areas, *Sci. Data*, 2017, vol. 4, pp. 122–170.
- Landau, S. and Everitt, B.S., *A Handbook of Statistical Analyses using SPSS*, CRC, 2004.
- LaRue, M.A. and Nielsen, C. K., Modelling potential dispersal corridors for cougars in midwestern North America using least-cost path methods, *Ecol Model*, 2008, vol. 212, pp. 372–381.
- Levi, M.R. and Rasmussen, C., Covariate selection with iterative principal component analysis for predicting physical soil properties, *Geoderma*, 2014, vol. 219, pp. 46–57.
- Li, X., Xiao, T., and Wang, L., Study on climate seasonal division and variation characteristics in Sichuan region in the past 44 years, *J. Chengdu Univ. Inf. Technol.*, 2007, vol. 22, pp. 531–538.

- Li, R.Q., Xu, M., Wong, M.H.G., Qiu, S., Sheng, Q.K., Li, X.H., et al., Climate change-induced decline in bamboo habitats and species diversity: implications for giant panda conservation, *Diversity Distrib.*, 2015, vol. 21, pp. 379–391.
- Li, D., Wintle, N. J. P., Zhang, G., Wang, C., Luo, B., Martin-Wintle, M.S., et al., Analyzing the past to understand the future: natural mating yields better reproductive rates than artificial insemination in the giant panda, *Biol. Conserv.*, 2017, vol. 216, pp. 10–17.
- Liu, C., Berry, P., Tp, Pearson, R., Selecting thresholds of occurrence in the prediction of species distributions, *Ecography*, 2010, vol. 28, pp. 385–393.
- Liu, B., Jiao, Z., Ma, J., Gao, X., Xiao, J., Hayat, M.A., et al., Modelling the potential distribution of arbovirus vector *Aedes aegypti* under current and future climate scenarios in Taiwan, China, *Pest Manage. Sci.*, 2019, vol. 75, pp. 3076–3083.
- Lu, Y., Liu, Y., Wang, Y., Liu, Y., Wang, A., Impact of climate change on bioclimatic types in a mountain area—a case from Sichuan Province, *Acta Ecol. Sin.*, 2015, vol. 35, pp. 7052–7060.
- Luo, M., Wang, H., Lyu, Z., Evaluating the performance of species distribution models Biomod2 and MaxEnt using the giant panda distribution data, *J. Appl. Ecol.*, 2017, vol. 28, pp. 4001–4006.
- Ma, K., Liu, D., Wei, R., Zhang, G., Xie, H., Huang, Y., et al., Giant panda reintroduction: factors affecting public support, *Biodivers. Conserv.*, 2016, vol. 25, pp. 2987–3004.
- Martin-Wintle, M.S., Kersey, D.C., Wintle, N.J.P., Aitken-Palmer, C., Owen, M.A., and Swaisgood, R.R., Comprehensive breeding techniques for the giant panda, *Adv. Exp. Med. Biol.*, 2019, vol. 1200, pp. 275–308.
- Merckx, B., Steyaert, M., Vanreusel, A., Vincx, M., and Vanaverbeke, J., Null models reveal preferential sampling, spatial autocorrelation and overfitting in habitat suitability modelling, *Ecol. Model.*, 2011, vol. 222, pp. 588–597.
- Minasny, B., Rossiter, D., Lark, M., Heuvelink, G., and Brus, D., Maps and models are never valid, but they can be evaluated, 2017.
- Miquelle, D.G., Rozhnov, V.V., Ermoshin, V., Murzin, A.A., Nikolaev, I.G., Hernandez-Blanco, J.A., et al., Identifying ecological corridors for Amur tigers (*Panthera tigris altaica*) and Amur leopards (*Panthera pardus orientalis*), *Integr. Zool.*, 2015, vol. 10, pp. 389–402.
- Peters, W., Hebblewhite, M., Cavedon, M., Pedrotti, L., Mustoni, A., Zibordi, F., et al., Resource selection and connectivity reveal conservation challenges for reintroduced brown bears in the Italian Alps, *Biol. Conserv.*, 2015, vol. 186, pp. 123–133.
- Phillips, S.J., Anderson, R.P., Dudik, M., Schapire, R.E., and Blair, M.E., Opening the black box: an open-source release of Maxent, *Ecography*, 2017, vol. 40, pp. 887–893.
- Qi, D., Impact of Minshan Mountain mineral resources development on giant panda population and its habitat, in *WWF Giant Panda Conservation Research Report*, Fan, Z., Ed., World Wide Fund for Nature or World Wildlife Fund, Beijing, 2009, vol. 1, pp. 1–42.
- Rayfield, B., Fortin, M.-J., and Fall, A., The sensitivity of least-cost habitat graphs to relative cost surface values, *Landscape Ecol.*, 2010, vol. 25, pp. 519–532.
- Shao, J., Li, Y., and Ni, J., The characteristics of temperature variability with terrain, latitude and longitude in Sichuan-Chongqing Region, *J. Geogr. Sci.*, 2012, vol. 22, pp. 223–244.
- Sichuan Forestry Department, *The 4th Survey Report on Giant Panda in Sichuan Province*, Chengdu: Sichuan Science and Technology Press, 2015.
- Swaisgood, R.R., Wei, F., McShea, W.J., Wildt, D.E., Kouba, A.J., and Zhang, Z., Can science save the giant panda (*Ailuropoda melanoleuca*)? Unifying science and policy in an adaptive management paradigm, *Integr. Zool.*, 2011, vol. 6, pp. 290–296.
- Tosi, G., Chirichella, R., Zibordi, F., Mustoni, A., Giovannini, R., Groff, C., et al., Brown bear reintroduction in the Southern Alps: to what extent are expectations being met?, *J. Nat. Conserv.*, 2015, vol. 26, pp. 9–19.
- van Gils, H., Westinga, E., Carafa, M., Antonucci, A., and Ciaschetti, G., Where the bears roam in Majella National Park, Italy, *J. Nat. Conserv.*, 2014, vol. 22, pp. 23–34.
- Vaughan, I., Ormerod, S., Methodological insights: increasing the value of principal components analysis for simplifying ecological data: a case study with rivers and river birds, *J. Appl. Ecol.*, 2005, vol. 42, pp. 487–497.
- Wang, F., McShea, W., Wang, D., Li, S., Zhao, Q., Wang, H., et al., Evaluating landscape options for corridor restoration between giant panda reserves, *PLoS One*, 2014, vol. 9, e105086.
- Wang, L. Wilderness and release of the giant panda “Taotao” successfully arrested. http://www.chengdu.gov.cn/chengdu/home/2017-12/31/content_f7ec0f3ac-48848c088cd9e9182660176.shtml?y7bRbP=KaR1kqrlplplplp-WfYX325D1dPoMb.Z9asB3yBccEqqlA. Accessed December 5, 2019.
- Wang, Y. and Deng, L., Quantitative structural analysis of land use and study on land sustainable utilization in Sichuan Province, *J. Sichuan Agricult. Univ.*, 2006, vol. 24, pp. 194–200.
- Xinhuanet. Latest data release of 2019 Giant Panda: the number of captive giant pandas in the world reaches 600. http://www.xinhuanet.com//2019-11/12/c_1125223631.htm. Accessed December 5, 2019.
- Xiong, M., *Study on Habitat Fragmentation Evaluation and Corridor Construction of Wolong Giant Panda*, Chengdu University of Technology, 2018.
- Yang, Y., Environmental impact assessment index system of mineral resources planning in sichuan province, *Metal Mine*, 2009, vol. 39, pp. 144–146.
- Yi, T. and Jiang, X., Staple food bamboo species of giant panda and their biodiversity, *J. Sichuan For. Sci. Technol.*, 2010, vol. 31, pp. 1–20.
- Yin, K., Xie, Y., and Ning, W.U., Corridor connecting giant panda habitats from north to south in the Min Mountains, Sichuan, China, *Integr. Zool.*, 2010, vol. 1, pp. 170–178.
- Yu, Z. The panda Zhang Meng, who was released in Yaan last October, ran to Mianming. <http://sc.sina.com.cn/news/m/2017-06-27/detail-ifyhmtk7838382.shtml>. Accessed December 5, 2019.
- Zhang, Z. and Wei, F., *Theory and Practice of Ex Situ Conservation of Giant Pandas*, Beijing:Beijing Science Press, 2006.
- Zhang, M., Huang, Y., Li, D., Zhang, H., Zhou, S., Huang, J., et al., Habitat selection of the pre-released giant panda in Wolong Nature Reserve, *Acta Ecol. Sin.*, 2013, vol. 33, pp. 6014–6020.
- Zhao, Z., Luan, X., Chen, J., Ye, J., Li, J., Zhang, C., et al., Improving management of Giant Panda National Park by quantitative assessment of ecological management, *Acta Ecol. Sin.*, 2019, vol. 39, pp. 3885–3894.
- Zhou, X., *Sichuan Expressway Management and Research*, Chengdu: Sichuan University, 2002.
- Ziółkowska, E., Ostapowicz, K., Kuemmerle, T., Perzanowski, K., Radeloff, V. C., and Kozak, J., Potential habitat connectivity of European bison (*Bison bonasus*) in the Carpathians, *Biol. Conserv.*, 2012, vol. 146, pp. 188–196.