Environ Monit Assess (2019) 191: 4
https://doi.org/10.1007/s10661-018-7130-4

Habitat selection and prediction of the spatial distribution of the Chinese horseshoe bat (*R. sinicus*) in the Wuling Mountains

Liang Liang · Xiang Luo · Zhixiao Liu · Jiahui Wang · Ting Huang · Erzhu Li

Received: 12 June 2018 / Accepted: 23 November 2018 / Published online: 5 December 2018

Abstract Habitat selection by the Chinese horseshoe bat (*Rhinolophus sinicus*) in the Wuling Mountains was studied in this paper. Global positioning system (GPS), remote sensing (RS) and geographic information system (GIS) technologies were used to obtain ground survey data and analyse the habitat factors driving the distribution of *R. sinicus*. Based on these basic data, a binary logistic regression method was used to establish habitat selection models of *R. sinicus*. Then, the corrected Akaike information criterion (AIC_C) was used to screen an optimal model, and the Hosmer-Lemeshow test indicated that the optimal model has suitable goodness of fit. Finally, the optimal model was used to predict the spatial distribution of *R. sinicus* in the Wuling Mountains. Verification analysis showed that the overall accuracy of the model was 72.7% and that the area under the curve (AUC) value was 0.947, which indicated that the model was effective for predicting suitable habitat for *R. sinicus*. The model results also showed that the main factors that influenced habitat selection were slope, annual mean temperature and distances from roads, rivers and residential land. *R. sinicus* preferred areas far from roads and residential land and areas near rivers. Generally, higher values of slope and annual mean temperature were associated with a greater likelihood of *R. sinicus* presence. Therefore, the protection of the water bodies surrounding *R. sinicus* habitats and fully addressing the impacts of human activities on *R. sinicus* habitats are recommended to protect the survival and reproduction of the population.

Keywords Habitat selection · Spatial distribution prediction · *Rhinolophus sinicus* · Logistic regression model

L. Liang (✉) · X. Luo · J. Wang · T. Huang · E. Li
School of Geography, Geomatics and Planning, Jiangsu Normal University, Xuzhou 221116 Jiangsu, China
e-mail: liangliang198119@163.com

L. Liang
School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210008 Jiangsu, China

Z. Liu (✉)
College of Biology and Environment Science, Jishou University, Jishou 416000 Hunan, China
e-mail: zxliu1965@163.com

Introduction

Research on the mechanisms of habitat selection and the prediction of the spatial distribution of species can objectively reflect the inherent regularity of species' habitat choice. This type of research is of great relevance for the protection and maintenance of biological diversity, especially for the protection of endangered species (Wang 2009; Wang and Li 2008).

The habitat selection mechanisms of animals are susceptible to various factors (Atuo and O'Connell 2017; Atuo et al. 2016; Zhang et al. 2014; Li et al. 2016). If animals' habitat selection is studied using a single data source, the habitat characteristics cannot be described in detail, and certain key factors are likely to be neglected (Jiang 2007). Thus, using multisource data to study habitat selection can comprehensively reflect the characteristics of animals' habitat selection. Species distribution models (SDMs) are important measures of species' habitat selection. SDMs primarily use data on the distribution of species and the corresponding environmental data to estimate the niche of a species using various algorithms. The results are then applied to the study area, and the distribution of the target species is estimated (Li et al. 2013; Xu et al. 2015). Many scholars select only environmental factors for modelling and ignore anthropogenic disturbance factors in SDMs (Fu et al. 2011; Vitousek et al. 1997). However, anthropogenic disturbance is an important factor in habitat selection by animals (Dudgeon et al. 2006). We sought to address these issues in the development and validation of SDMs by creating occupancy models and mapping the distribution of the Chinese horseshoe bat (*Rhinolophus sinicus*).

R. sinicus is a typical cave-dwelling species; it remains in the cave during the day but flies outside the cave to prey on insects (mostly agricultural and forestry pests, mainly Lepidoptera and Coleoptera species) at night. *R. sinicus* is distributed in the Himalayas, China and Vietnam (Xie et al. 2017). In China, *R. sinicus* is mainly distributed in the area south of the Yangtze River (Wang 2003). According to the 2016 World Conservation Union Red List (IUCN 2016), 29.1% of the population of *Rhinolophus* showed a decreasing trend due to habitat fragmentation and anthropogenic disturbance. There has been extensive research on the cytogenetics (Xu 2012), predation, predation behaviour (Miková et al. 2013), habitat selection (Yu et al. 2015) and echolocation (Zhang et al. 2008) of *R. sinicus*. However, research on habitat selection mechanisms, the factors affecting habitat selection and spatial distribution prediction for *R. sinicus* remains lacking in an important *R. sinicus* distribution area—the Wuling Mountains.

The existing research on the dependence of bats on caves was mainly conducted by investigating caves to obtain point source information, including the geographical locations of the caves, altitude and vegetation status (Yu et al. 2015), which is time consuming, labourious and inefficient. With the development of spatial information technology, animal habitats have been monitored and mapped synchronously with "3S" technology, which includes global positioning system (GPS), remote sensing (RS) and geographic information system (GIS). "3S" technology has the advantages of covering a large area and being fast and highly efficient, and it can also obtain surface source information (Chen et al. 2015). Therefore, using "3S" technology for data acquisition is not only time saving and labour saving but also scientific and objective.

A generalised linear model (GLM) is a type of SDM that represents the generalisation of multiple linear regression models (Xu et al. 2015). Among multiple linear regression models, logistic models are the most well known (Deng 2010) and have been widely used to study habitat selection by wildlife (Atuo et al. 2016; Fu et al. 2011; Ji et al. 2007). In the present study, we created and validated a binary logistic regression model based on "3S" technology from the perspective of multisource data (Sukumal et al. 2010; Li et al. 2012), which can be used to identify high-priority areas for conservation and provide important theoretical guidance for the conservation of species and the maintenance of species diversity. Specifically, our objectives were to (1) determine how habitat factors influence habitat selection by *R. sinicus*, (2) determine the interference effects of anthropogenic disturbance factors on habitat selection by *R. sinicus* and (3) analyse the primary factors influencing species presence and the habitat selection characteristics of *R. sinicus* to predict the spatial distribution of this species.

Materials and methods

Study area

The study area spans the main distribution area of *R. sinicus* (25° 52′ to 31° 24′ N and 107° 04′ to 112° 02′ E) and encompasses four administrative provinces and cities (Hunan, Hubei, Chongqing and Guizhou; Fig. 1). The area has a total of approximately 0.17 million km^2 with an average elevation of 700 m and an average slope of 15°. The climate in this region is a transition zone from subtropical to warm temperate, the annual average temperature is 13.5 to 17.0 °C, and the total annual precipitation is between 1100 and 1600 mm. The forest cover types in this area are mainly coniferous and broad-leaved mixed forest and evergreen coniferous forest, and the average forest coverage is 56%, which is much higher than the national average of 20.36% (You et al. 2013). The carbonate karst caves in this region are well developed. The study area is one of the main distribution areas of bats in China (Li et al. 2005).

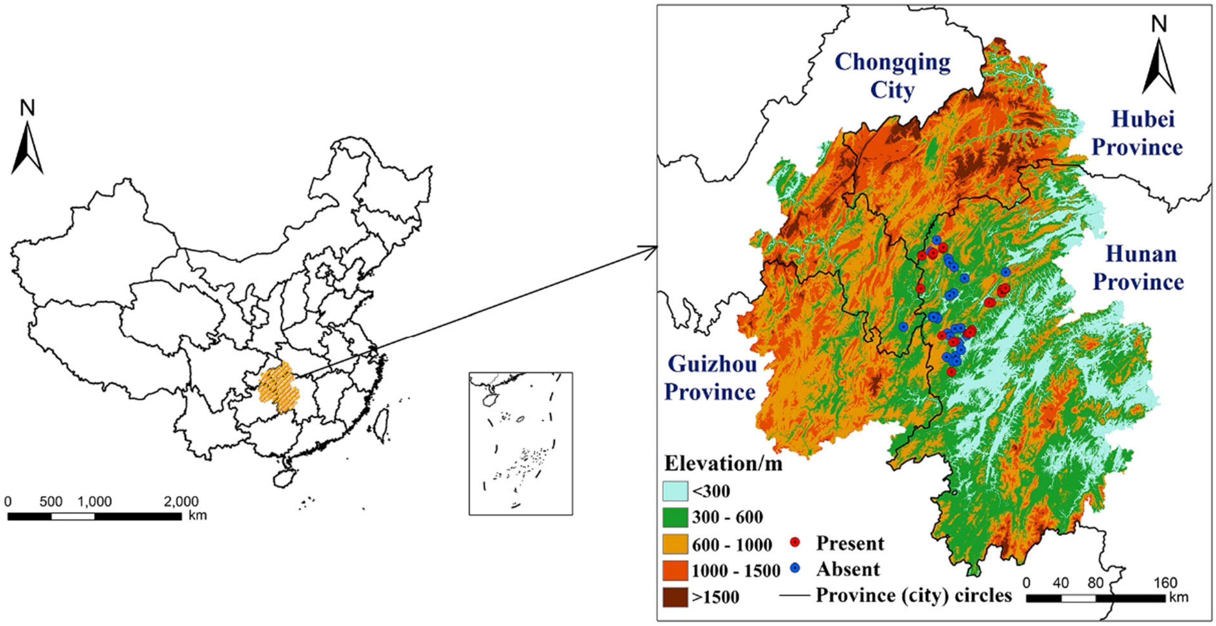

Fig. 1 Overview of the study area. The background is the elevation of the study area. The points are survey points from 2015 to 2017

Sample collection

Several data sources were used in the binary logistic regression model. From January 2015 to April 2017, field surveys of caves were conducted in the Wuling Mountains. *R. sinicus* is a typical cave-dwelling species and is highly dependent on caves. Therefore, it is important to know the distribution of caves in the study area. The random point method (Mcconnell et al. 2008) was used to determine the area to be investigated, and the caves in the area were investigated by visiting local residents and guides.

During the study, we took the time during which *R. sinicus* rests in the cave as the observation period (9:00–17:00). Caves were determined to represent *R. sinicus* habitat based on observations and recordings of excrement, food trails, naturally dead individuals and living individuals of *R. sinicus* in the caves. At each sampling locality, we observed and recorded habitat information such as topography (altitude, slope and aspect), vegetation status, cave characteristics (the size of the hole and the length of the cave, etc.) and human disturbance (farmlands, roads and residential land) inside and around the cave and recorded the geographical location of each cave using GPS (Garmin International, Inc., Olathe, KS, USA). ArcGIS10.2 (ESRI, Redlands, CA, USA) was used to generate the survey point vector data. In this study, 55 caves were investigated. Twenty-nine caves were identified as habitats of *R. sinicus* (Fig. 1).

To identify *R. sinicus*, bats were captured with mist nets as they left their day roosts. For the captured bat individuals, we recorded the sex, measured morphological characters and identified them to species.

Selection and normalisation of factors

A number of abiotic, biotic and anthropogenic factors can influence the quality and availability of species' habitat (Ouyang et al. 2011; Fan et al. 2014). To explore *R. sinicus* habitat selection mechanisms and describe *R. sinicus* habitat, ten factors were selected from among four categories of habitat factors: vegetation factors, topographic feature factors, meteorological factors and anthropogenic disturbance factors.

Vegetation factors Vegetation coverage (VFC) is a good indicator of vegetation status (Liu et al. 2011; Uhrin et al. 2017). RS data can be used to obtain a thematic map of vegetation indices (Liang et al. 2015, 2016, 2017). In the present study, the thematic map mainly contained normalised difference vegetation index (NDVI) data, with outliers removed, based on Landsat 8 OLI data from 2015 to 2017, which were obtained from the Geospatial Data Cloud. The land use data were available as global surface coverage data from GlobeLand30 and were downloaded from the Global Land Cover website. Based on the NDVI data and land use data, a VFC map of the study area was retrieved using the dimidiate pixel model (Zribia et al. 2003; Wu et al. 2017; Zhou et al. 2017).

Topographic feature factors Topographic feature factors include elevation, slope and aspect (Wang 2009; Uhrin et al. 2017). Based on digital elevation model (DEM) data, slopes and aspects were calculated by the GIS spatial analysis module. The DEM data were downloaded from the Geospatial Data Cloud at 30-m spatial resolution from the GDEMDEM data.

Meteorological factors The meteorological factors comprised annual mean temperature and annual precipitation. After pretreatment, the annual mean temperature and annual precipitation were extracted. The meteorological data were downloaded from the global meteorological data set of the WorldClim Version 2 of the World Meteorological Database (http://www.worldclim.org/).

Anthropogenic disturbance factors The anthropogenic disturbance factors were the nearest distances from roads, rivers, residential land and farmland (Kelly et al. 2016; Gerald and Markus 2009; Agosta et al. 2005; Lumsden et al. 2002). The Euclidean distance calculation method (Montechiesi et al. 2016) was used to determine the nearest distances from roads, rivers, farmland and residential land at each point (30 × 30-m grid) in the study area. In this study, river, farmland and residential land data were extracted from the land use data. The land use data were available as global surface coverage data from GlobeLand30 and were downloaded from the Global Land Cover website. The road vector data were downloaded from the DIVA-GIS website.

All statistical analyses of the data were performed using SPSS18.0 (IBM, New York City, New York, USA). Because the units of each factor were inconsistent, a unified dimension of each factor was required for modelling, and all habitat factors were normalised to remove the effects of dimension.

Habitat factor analysis before modelling

To analyse the effect of each factor on habitat selection by *R. sinicus*, the importance of the factors was explored by difference analysis (comparing the difference of each factor between the presence and absence of *R. sinicus*). In this study, the Kolmogorov-Smirnov test (K-S test) was used to test the normality of the factors (Wang 2009; Fu et al. 2011). For normally distributed factors, an independent sample *t* test was used to analyse the significance of the difference in each factor between the presence and absence of *R. sinicus* (Fu et al. 2011). For the factors with non-normal distributions, the Mann-Whitney *U* test was used to analyse the significance of the differences of the original data (Wang 2009; Fu et al. 2011).

To test the independence of the above habitat factors and avoid the influence of collinearity among the factors, it was necessary to analyse the correlation of each factor before modelling. During the correlation analysis, one of the factors was removed according to the importance of ecological significance when the correlation coefficient between two factors was greater than 0.75 (Fu et al. 2011).

Modelling

The binary logistic regression model has been widely used to study and predict habitat selection by species because it can predict the probability of occurrence accurately (Fu et al. 2011; Ji et al. 2007). In this study, we extracted the values of habitat factors corresponding to each surveyed spot from the habitat factor thematic map described in the "Selection and normalisation of factors" section. Then, the extracted habitat factors' values were entered as independent variables in different combinations (10 factors, 1023 combinations); *R. sinicus* presence/absence was entered as a dependent variable, and all possible models were implemented. The basic formula of the binary logistic regression model was as follows (Ji et al. 2007; Li et al. 2012):

$$\mathrm{Logit}\frac{P(Y=1)}{1-P(Y=1)} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_i x_i \quad (1)$$

where α is constant of the equation, x_i is a predictor factor and β_i is the coefficient of the predictor factor.

To screen the optimal models, the corrected Akaike information criterion (AIC_C), a modification of Akaike's information criterion (AIC), was used to determine significant differences among different models and comprehensively weigh the relationship between the applicability of a model and the number of parameters (Hurvich et al. 1990; Akaike 1973, 1974; Atuo and O'Connell 2017; Atuo et al. 2016; Shu et al. 2010). W_i

is the weight of each model, which can represent the relative importance of the model. AIC_C and W_i were determined using the following formulas (Shu et al. 2010; Burnham and Anderson 2002):

$$AIC_C = -2ln\left[L\left(D|\hat{\theta}\right)\right] + 2\left(\frac{n}{n-m-1}\right) \quad (2)$$

$$\Delta AIC_C = AIC_{Ci} - AIC_{Cmin} \quad (3)$$

$$W_i = \frac{\exp\left(-\frac{\Delta AIC_{Ci}}{2}\right)}{\sum_{r=1}^{R} \exp\left(-\frac{\Delta AIC_{Ci}}{2}\right)} \quad (4)$$

where $ln\left[L\left(D|\hat{\theta}\right)\right]$ is the maximum log likelihood, m is the number of model parameters including the number of factors and the intercept and n is the sample number.

In this study, AIC_C and W_i were used to select the most appropriate regression model and determine the primary influencing factors. First, the model with the lowest AIC_C value was selected as the optimal model, and the factors contained in the model were considered the main influencing factors. Second, any model within $\Delta AIC_C < 2.0$ of the optimal model was considered acceptable (Burnham and Anderson 2002). Third, the weighted coefficients of the main influencing factors in these models were summed to obtain the final coefficients of main influencing factors and the logistic regression model was established.

In logistic analysis, the Hosmer-Lemeshow test is typically used to verify the fitting effect of the model. Therefore, the Hosmer-Lemeshow test was used to evaluate the fitting degree of the model in this study. The Hosmer-Lemeshow test reflects the difference between the expected frequency and the observed frequency. If the difference is not significant, the chi-square value is small and the significance value is large (Hosmer and Lemeshow 2000; Gao et al. 2017).

Habitat distribution prediction and accuracy evaluation

Using the thematic map of the main influencing factors as the input variables in the established model, the spatial distribution *R. sinicus* was predicted in the Wuling Mountains. Then, we used the 55 surveyed sites to evaluate the prediction results. First, we adopted a threshold method to distinguish suitable habitats and unsuitable habitats for *R. sinicus* and then calculate the overall prediction accuracy (Qin et al. 2017; Abdukerim et al. 2016). In addition, the receiver operating characteristic (ROC) curve was used to evaluate the model. The ROC curve takes each value of a predicted result as a possible judgement threshold, thereby calculating the corresponding sensitivity and specificity. Because this method combines the advantages of the sensitivity and specificity indices and is simple, the ROC curve has been widely used in species distribution modelling (Wang et al. 2007; Qin et al. 2017).

Results and analysis

The distribution characteristics of habitat factors in the Wuling Mountains

The VFC in the Wuling Mountains area ranged from 0 to 100%, and the average value was approximately 83%. Regions with VFC greater than 90% accounted for 93.7% of the total study area (Fig. 2a), which indicated good vegetation growth in the Wuling Mountains. You et al. (2013) showed that the forest coverage rate reached 56% in the Wuling Mountains, which was much higher than the national average of 20.36%.

Regarding the topographic feature factors, the average elevation of the Wuling Mountains was approximately 702 m and the average slope was 16°. As the Wuling Mountains lie in the western and northern regions of the study area, the elevation and slope of the western and northern regions were greater than those of the eastern and southern regions (Fig. 2b, c). Due to the large number of ranges in the study area, the number of grids in each of the four aspects of the study area was similar (Fig. 2d).

The climate in this region was a transition zone from subtropical to warm temperate, the annual mean temperature was 13.5 to 17.0 °C, and the annual mean temperature in the north was lower than that in the south (Fig. 3a). The total annual precipitation was between 1100 and 1600 mm, and the annual precipitation in the east was greater than that in the west (Fig. 3b).

The anthropogenic disturbance factors had a negative effect on habitat loss and fragmentation, and we considered the effects of anthropogenic disturbance on habitat selection. The Euclidean distance calculation method was used to obtain the nearest distances from roads,

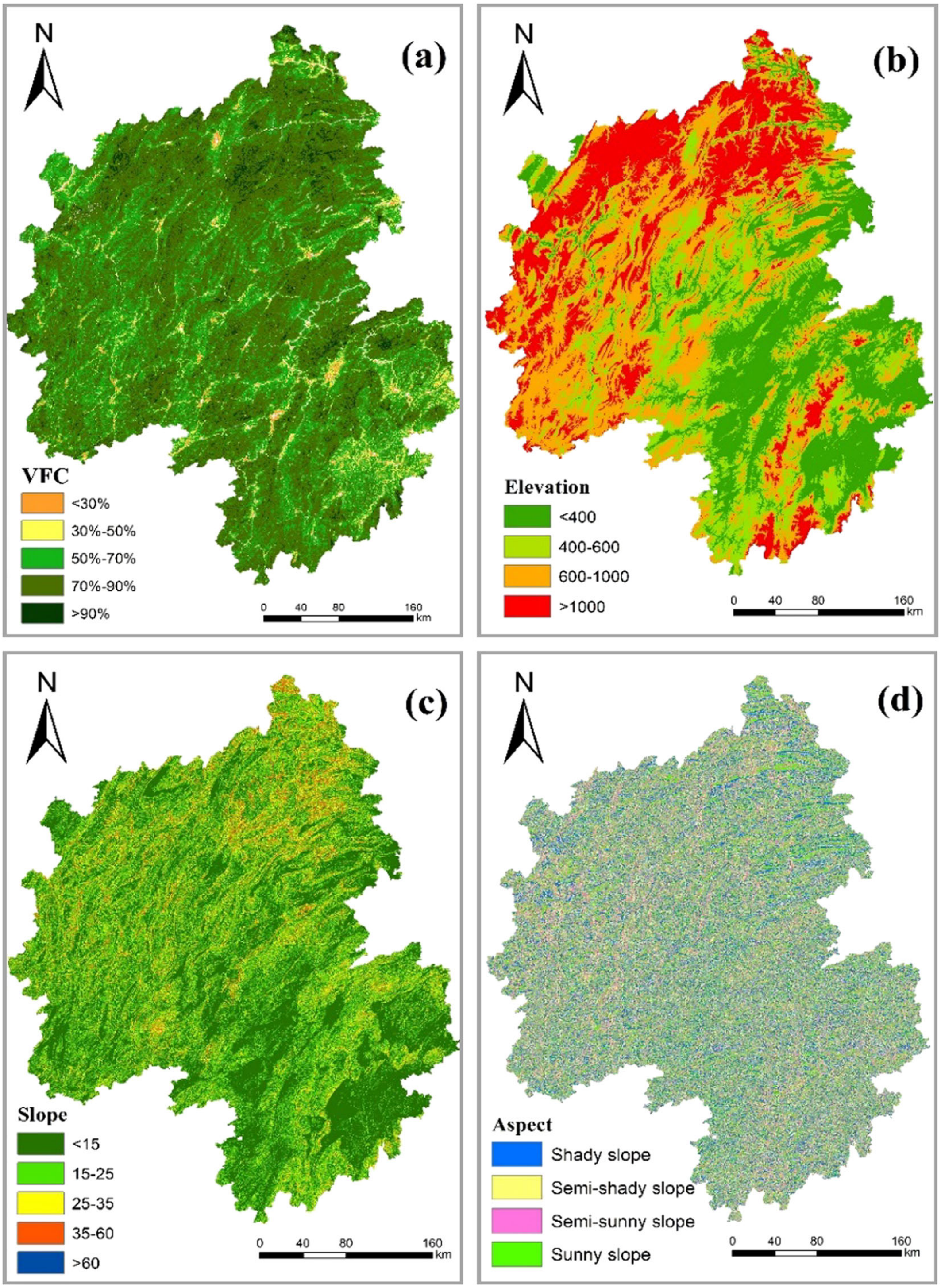

Fig. 2 The vegetation coverage and topographic features in the study area. **a** Vegetation coverage. **b** Elevation. **c** Slope. **d** Aspect

rivers, farmland and residential land at each point (30 × 30-m grid) of the study area (Fig. 4). Among all survey sites, the minimum distance from roads was 242 m, and the maximum distance was 33,062 m; the minimum distance from rivers was 518 m, and the maximum distance was 9332 m; the minimum distance from

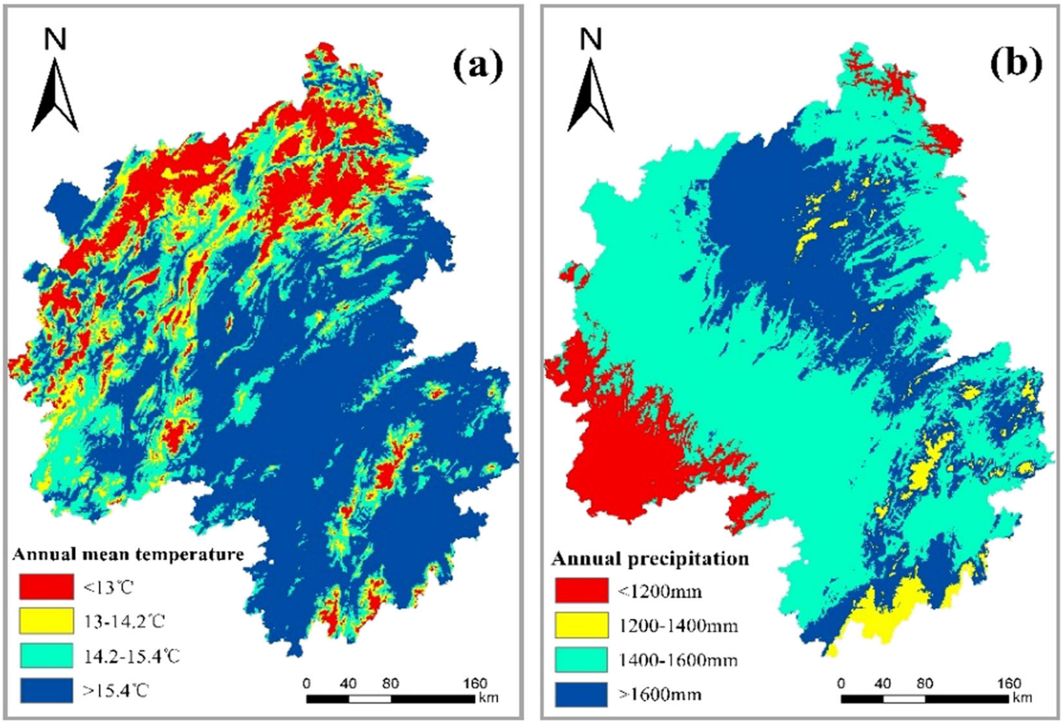

Fig. 3 The annual mean temperature (**a**) and annual precipitation (**b**) in the Wuling Mountains area

residential land was 189 m, and the maximum distance was 8561 m; and the minimum distance from farmland was 15 m, and the maximum distance was 1570 m.

The importance and independence of habitat factors

This paper examined the factor differences between the areas where *R. sinicus* was present and those where it was absent. The K-S test showed that the factors such as slope, elevation, vegetation coverage, annual mean temperature, annual precipitation and nearest distances from roads and rivers exhibited normal distributions (thus, they were analysed using an independent sample *t* test), whereas the nearest distances from residential land and farmland did not (thus, they were analysed using a Mann-Whitney *U* test). The results indicated that *R. sinicus* preferred areas with higher elevation, steeper slopes and higher VFC. At the same time, these areas are close to rivers; far away from roads, residential land and farmland; and have suitable temperatures (approximately 15.6 °C) and abundant precipitation (approximately 1400 mm). In addition, the results showed that the differences in slope, elevation, annual mean temperature, annual precipitation and nearest distances from rivers, roads, residential land and farmland were significant between *R. sinicus* presence areas and absence areas, whereas VFC was not significantly different between areas (Table 1). This latter finding is likely due to the high total VFC in the Wuling Mountains. That is, although vegetation coverage is an important factor affecting the distribution of *R. sinicus* in other places (Froidevaux et al. 2016), the high overall VFC makes it no longer a limiting factor in the Wuling Mountains.

To test the independence of each factor, the correlations of habitat factors were analysed. The result showed that the correlation coefficients between each factor were less than 0.75 (Table 2), indicating that there was no multicollinearity between the factors; therefore, the model will be established by using all of the factors mentioned in the "Materials and methods" section in the subsequent process.

Habitat selection model of *R. sinicus*

The habitat factors identified by the correlation and difference analyses were used to establish the logistic regression models, and a total of 1023 models were established. AIC_C and W_i were calculated for each model, followed by the value of ΔAIC_C. Among them, the model with the smallest AIC_C was selected as the optimal

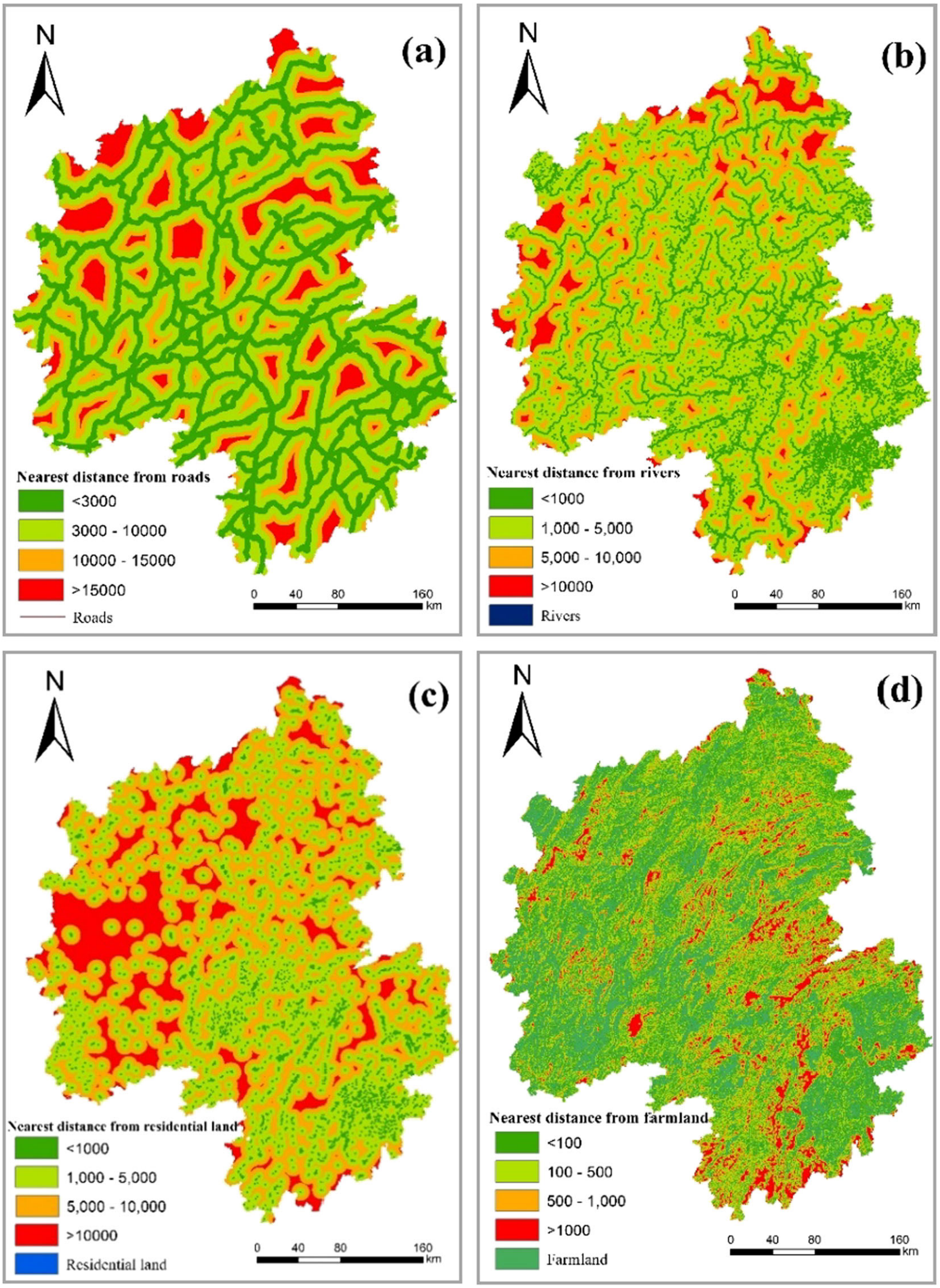

Fig. 4 The anthropogenic disturbance factors. Nearest distance from **a** roads, **b** rivers, **c** residential land and **d** farmland

model, which included the factors nearest distances from rivers, roads and residential land; annual mean temperature; and slope, indicating that these factors were the main factors influencing habitat selection in *R. sinicus*. Then, all of the remaining models with $\Delta AIC_C < 2.0$ were filtered out, and the weighted sums of the main factor

Table 1 Comparisons of the factors between presence and absence areas of *R. sinicus*

Factor	Present	Absent	K-S test *Z*-value	Independent sample *t* test *t*-value	Mann-Whitney *U* test *U*-value
Elevation	590.690 ± 260.972	499.385 ± 198.572	1.043	0.631*	–
Slope	20.630 ± 9.523	12.236 ± 10.957	0.979	0.194*	–
Vegetation coverage	0.788 ± 0.063	0.774 ± 0.047	0.951	0.365	–
Nearest distance from rivers	3028.022 ± 1887.083	3914.003 ± 1904.553	0.577	0.289*	–
Nearest distance from roads	5806.966 ± 4105.416	2828.615 ± 2994.518	1.215	9.866*	–
Nearest distance from residential land	3744.573 ± 2169.381	2136.203 ± 1750.659	2.009	–	163*
Nearest distance from farmland	313.049 ± 342.814	124.174 ± 191.958	1.662	–	211.5*
Annual mean temperature	15.61 ± 1.17	16.39 ± 0.93	1.282	1.160*	–
Annual precipitation	1418.97 ± 53.44	1376 ± 47.43	1.258	1.074*	–

*Difference is significant at the 0.05 level (two-tailed test). Because aspect is a qualitative factor (including sunny slope, semi-sunny slope, semi-shady slope and shady slope) and the difference analysis generally considers quantitative factors, this paper did not carry out a difference analysis of aspect

coefficients were calculated (Table 3). Finally, the probability model of *R. sinicus* was obtained:

$$\mathrm{Logit}\frac{P}{1-P} = 0.013 + 0.13 \times \mathrm{Slope} + 0.171 \times \mathrm{Road} - 0.774 \times \mathrm{River} + 0.352 \times \mathrm{Residents} - 0.496 \times \mathrm{BIO1} \quad (5)$$

where Slope is the normalised slope, Road is the normalised nearest distance from roads, River is the normalised nearest distance from rivers, Residents is the normalised nearest distance from residential land and BIO_1 is the normalised annual mean temperature.

The Hosmer-Lemeshow test was used to evaluate model fit. Our calculations showed that the expected values were close to the observed values (Table 4). We set 0.05 as the significance level such that the critical value was CHIINV (0.05, 7) = 14.067. As shown in Table 4, the Hosmer-Lemeshow test yielded a chi-square value of 2.329, which is less than the critical value, and the statistic was not significant ($P = 0.939$).

Table 2 Pearson correlation coefficients between driving factors

Habitat factors	1	2	3	4	5	6	7	8	9	10
1	1									
2	− 0.015	1								
3	0.049	− 0.12	1							
4	− 0.061	0.595	0.033	1						
5	0.100	0.229	0.589	− 0.030	1					
6	− 0.031	0.442	0.126	0.205	0.419	1				
7	− 0.014	0.209	0.261	0.220	0.462	0.355	1			
8	0.002	0.182	0.417	0.265	0.301	0.108	0.383	1		
9	0.040	− 0.673	− 0.150	− 0.450	− 0.409	− 0.417	− 0.273	− 0.239	1	
10	− 0.056	0.696	0.159	0.327	0.463	0.272	0.269	0.201	− 0.681	1

1, aspect; 2, elevation; 3, slope; 4, vegetation coverage; 5, nearest distance from roads; 6, nearest distance from rivers; 7, nearest distance from residential land; 8, nearest distance from farmland; 9, annual mean temperature; 10, annual precipitation

Table 3 Weighted average sums of the factor coefficients in the models with $\Delta AIC_C < 2.0$

Model	Model coefficients					W_i	Weighted average sum of coefficients				
	Slope	Road	River	Residents	Tem		Slope	Road	River	Residents	Tem
Slope + Tem + River+ Residents	1.117		− 2.903	1.550	− 2.137	0.066	0.073		− 0.191	0.102	− 0.140
Residents + River + Road + Tem		1.394	− 3.291	1.416	− 1.855	0.063		0.089	− 0.209	0.090	− 0.118
Road + Residents + River + Slope + Tem	0.687	0.852	− 3.220	1.420	− 2.053	0.038	0.026	0.032	− 0.121	0.053	− 0.077
Slope + Farmland + River + Residents + Tem	0.937		− 3.187	1.489	− 2.275	0.033	0.031	0.05	− 0.104	0.048	− 0.074
Road + Farmland + River + Residents + Tem		1.151	− 3.419	1.341	− 1.984	0.044			− 0.150	0.059	− 0.087
Sum							0.13	0.171	− 0.774	0.352	− 0.496

The models represent the models with $\Delta AIC_C < 2.0$; the model coefficients were obtained from the binary logistic regression model; W_i is the weight of model; the weighted average sums of the coefficients were equal to the model coefficients multiplied by W_i, and then the product was summed

The results indicate that the model has suitable goodness of fit and can be used to predict the spatial distribution of *R. sinicus* in the Wuling Mountains.

Spatial distribution of *R. sinicus* in the Wuling Mountains

In this paper, the grid layers, including elevation and nearest distances from roads, rivers and residential land were plugged into the regression equation and the grid was calculated. Then, the probability map of spatial distribution of *R. sinicus* in the Wuling Mountains was obtained (Fig. 5). We considered areas with probabilities greater than 0.5 as areas of available habitat for *R. sinicus* and those with probabilities equal to or less than 0.5 as areas of non-available habitat according to Wang (2009). The results showed that the available habitat area was 83,408.7 km^2, accounting for 48.5% of the total study area, and was mainly distributed in the northern northwestern regions of the Wuling Mountains. The non-available habitat area was 88,534.9 km^2, accounting for 51.5% of the total study area, and was mainly distributed in the southern region of the Wuling Mountains. The available habitat for *R. sinicus* appeared

Table 4 Hosmer-Lemeshow goodness-of-fit test results for the logistic model

Group	Total	$Y = 1$		$Y = 0$	
		Observed	Expected	Observed	Expected
1	6	0	0.022	6	5.978
2	6	0	0.193	6	5.807
3	6	1	0.588	5	5.512
4	6	1	1.971	5	4.029
5	6	4	3.513	2	2.487
6	6	5	4.590	1	1.410
7	7	6	6.407	1	0.593
8	6	6	5.866	0	0.134
9	6	6	5.949	0	0.051
	Chi-square	DF	*P*		
	2.329	7	0.939		

Group represents the sample group; Total represents the number of samples in each group. $Y = 1$ denotes *R. sinicus* presence; $Y = 0$ denotes *R. sinicus* absence; Observed refers to the observed frequency; Expected refers to the expected frequency; Chi-square is the index value of the Hosmer-Lemeshow test; DF is degrees of freedom; *P* is statistical significance

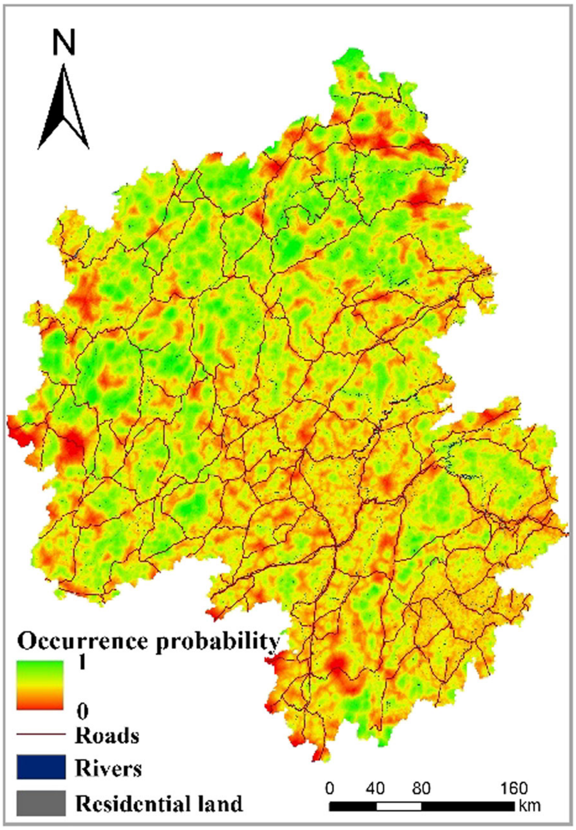

Fig. 5 Predicted spatial distribution of *R. sinicus*

to exhibit fragmentation in the Wuling Mountains, which poses a threat to the survival and reproduction of *R. sinicus*.

Prediction accuracy evaluation

In this study, surveyed sites were used to evaluate the prediction accuracy. According to Qin et al. (2017) and Abdukerim et al. (2016), we set a threshold of 0.2, and the probability of occurrence corresponding to each survey site was extracted. An occurrence probability value of the point corresponding to the presence of *R. sinicus* greater than 0.2 or one corresponding to the absence of *R. sinicus* less than 0.2 indicated that the prediction was correct; otherwise, the prediction was incorrect. Validation analysis showed that among all survey samples, 40 samples were correct; that is, the prediction overall accuracy was 72.7%, indicating that the model has good prediction performance.

To further evaluate the prediction accuracy, each value of a predicted result was used as a possible threshold to draw the ROC curve in this paper. Figure 6 shows the ROC curve obtained from the binary logistic regression model. The ROC curve is drawn with the false positive rate (1—specificity) as the abscissa and the true positive rate (sensitivity) as the ordinate, and the area under the curve (AUC) is taken as a measurement index of model prediction accuracy, which is obtained by integration (Xu et al. 2015; Wang et al. 2007). The values range from 0 to 1; the greater the value, the stronger the judgement of the model. In the study, the AUC value of the model prediction result was 0.947, which is higher than the value of 0.5 from a random distribution model, indicating that the prediction of the spatial distribution of *R. sinicus* in the Wuling Mountains is reliable.

Discussion

Habitat selection is a behavioural process in which species respond to and make decisions about available habitats. The factors that influence bat habitat selection include insect abundance (Goiti et al. 2015), forest type (Wang 2009), tree line (Kalda et al. 2015), microclimate conditions (Miková et al. 2013), elevation (Wang 2010; Bontadina 2002) and human disturbance factors (Kelly et al. 2016), indicating that bat habitat selection was complex. In this study, the binary logistic regression model was used to study the habitat selection mechanisms of *R. sinicus* and predict the spatial distribution of this species in the Wuling Mountains. The validation analysis indicated that the reliability of the model was

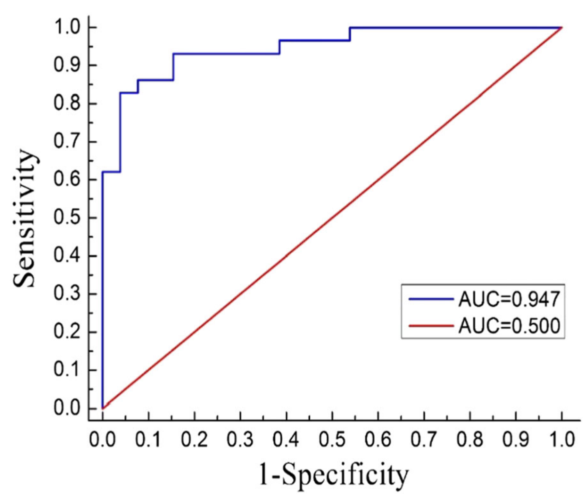

Fig. 6 Receiver operating characteristic curve

good, and the results can provide a scientific basis for relevant departments or decision makers.

The comprehensive model selection and difference analysis results revealed that *R. sinicus* preferred areas far from roads, residential land and farmland and areas near rivers. Human activities frequently occur near roads, residential land and farmland, which provided a certain constraint on the range of these activities. Gaisler et al. (1998) found a negative correlation between the number of bats and the distance from the centre of a city. Gerald and Markus (2009) found that roads could directly or indirectly reduce the number of animals and interfere with the activities of animals. Kelly et al. (2016) showed that the enhancement of agricultural production largely limited the ability of farmland to support other species, and agriculture was predicted to have a negative impact on bats and other species in the future. These results were consistent with those of the present study. The insect abundance was higher near rivers, which can provide a rich food source for *R. sinicus*. Additionally, rivers can provide the necessary water for *R. sinicus*. Studies have shown that rivers are generally considered to be important sources of food and drinking water for bats (Wang 2009), and this result was validated in the present study. At the same time, the spatial distribution prediction showed that there was fragmentation of the suitable habitat of *R. sinicus*, which might seriously affect the survival and reproduction of the species. Therefore, local government departments need to take conservation measures as soon as possible to maintain the local biodiversity according to the predicted distribution of suitable habitats.

Topographical features influenced the distribution of *R. sinicus*. Although the results of model selection indicated that elevation was not a major influencing factor, within a certain range, the higher the elevation, the greater the likelihood of *R. sinicus* presence. Almost all sites where *R. sinicus* was present were distributed in the mountainous areas above 500 m. Wang (2009) and Bontadina (2002) found that elevation was an important factor in habitat selection by *R. sinicus*, which is consistent with our findings. The aspect of a mountain has a substantial impact on light intensity. Although the model results showed that aspect was not an important predictor of *R. sinicus* habitat selection, 23 of the 29 sites with *R. sinicus* presence were located on semi-shady and semi-sunny slopes, indicating a preference of *R. sinicus* for such slopes. *R. sinicus* is a nocturnal species that consistently lives in dark environments and avoids light. Due to the impact of light intensity, sunny slopes are often dry and have high temperatures, which are not suitable conditions for *R. sinicus*. In addition, the temperatures of semi-shady and semi-sunny slopes are higher than those of shady slopes. Thus, the insect richness of semi-shady and semi-sunny slopes is higher than that of shady slopes because insects are very sensitive to temperature, which is beneficial for *R. sinicus* predation (Wang 2009).

Regarding the meteorological factors, annual mean temperature was an important variable affecting *R. sinicus* presence. Almost all of the sites where *R. sinicus* was present were distributed in an area with an annual mean temperature greater than 15 °C, and the suitable annual temperature was approximately 15.6 °C. However, a higher annual mean temperature did not correlate with a greater occurrence probability of *R. sinicus*. Although the results of model selection indicated that annual precipitation is not the main factor influencing habitat selection in the Wuling Mountains, *R. sinicus* prefers high-humidity areas, and the annual precipitation can affect the activity of *R. sinicus* by affecting the environmental humidity. Wang (2009) found that bat activity was affected by temperature and humidity, consistent with the results of the present study.

The difference in the response of bats to vegetation structure depends on the predation strategy, ecological characteristics, and wing shape. Bats with low manoeuvrability prefer open areas, while bats with high manoeuvrability prefer complex zones (Froidevaux et al. 2016). *R. sinicus* belongs to the Microchiroptera, which comprises small bats with high manoeuvrability; therefore, these bats prefer a complex zone with high VFC. However, the results showed that VFC did not influence habitat selection. The reason was likely that more than 93.6% of the total study area had VFC greater than 70%, which indicated that the entire study area has a complex vegetation structure. In that case, although VFC is important for *R. sinicus*, it is no longer a limiting factor affecting their distribution in the Wuling Mountains.

Conclusion

The study of habitat selection mechanisms is important for species interaction and for the survival, conservation and maintenance of species diversity. In this paper, "3S" technology and binary logistic regression methods were used to study the habitat selection mechanisms and predict the spatial distribution of *R. sinicus* in the Wuling Mountains. The major findings include the following:

1. Verification analysis of *R. sinicus*'s spatial distribution prediction showed that the overall accuracy was 72.7% and that the AUC value was 0.947, which indicated it was appropriate to use "3S" technology and binary logistic regression methods to study habitat selection by *R. sinicus* in the Wuling Mountains.
2. The results showed that *R. sinicus*'s available habitat accounts for approximately 48.5% of the Wuling Mountains, mainly in the north and northwest regions. It is noteworthy that the available habitat appeared to exhibit fragmentation in the Wuling Mountains, which poses a threat to the survival and reproduction of *R. sinicus*.
3. In the Wuling Mountains, the main factors that influenced habitat selection were slope, annual mean temperature and distances from roads, rivers and residential land. *R. sinicus* preferred areas far from roads and residential land and areas near rivers. Generally, high values of slope (greater than 15°) and suitable annual mean temperature (approximately 15.6 °C) were associated with a greater likelihood of *R. sinicus* presence.

Acknowledgements The authors particularly thank NASA for providing the basic data (including the Landsat series data and ASTER GDEM data) and the NGCC for providing the GlobeLand30 data.

Author contributions L. Liang and Z. X. Liu proposed the main idea and offered guidance to complete the work. L. Liang and X. Luo wrote the paper. Z. X. Liu provided ground survey data. X. Luo, J. H. Wang, T. Huang and E. Z. Li carried out the data processing and analysis.

Funding This research was supported by the National Natural Science Foundation of China (No. 31560130 and No. 41401473), the Natural Science Foundation of Jiangsu Province (BK20181474), the Open Fund of State Key Laboratory of Remote Sensing Science (OFSLRSS201804), the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) and the 2018 Jiangsu Province Graduate Research and Innovation Project (KYCX18_2157).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abdukerim, R., Zibibilla, G., Xu, Z. L., Nabi, E., & Abliz, A. (2016). Assessment of habitat suitability for Gazella subgutturosa in the Ebinur reserve. *Acta Ecologica Sinica, 36*, 4171–4177.

Agosta, S. J., Morton, D., Marsh, B. D., & Kuhn, K. M. (2005). Nightly, seasonal, and yearly patterns of bat activity at night roosts in the central Appalachians. *Journal of Mammalogy, 86*(6), 1210–1219.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. *International Symposium on Information Theory, 1*, 610–624.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Atuo, F. A., & O'Connell, T. J. (2017). Structural heterogeneity affects raptor assemblages and niche characteristics in mixed-grass ecosystems. *Ecosphere, 8*(9), e01907. https://doi.org/10.1002/ecs2.1907.

Atuo, F. A., Ivande, S. T., Wala, Z. J., Manu, S., & O'Connel, T. J. (2016). Current distribution, breeding population and habitat use of the globally threatened Grey-necked Picathartes Picathartes oreas in south-eastern Nigeria: a call for conservation action. *Ostrich, 87*(2), 101–112.

Bontadina, F (2002) Conservation ecology in the horseshoe bats Rhinolophus ferrumequinum and Rhinolopus hipposideros. Dissertation. University of Bern, Bern.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference*. New York: Springer ISBN 0-387-95364-7.

Chen, W. X., Yang, X. R., & Li, Z. Q. (2015). The analysis of forest landscape patterns in east Dongting Lake region based on 3S technology. *Chinese Agricultural Science Bulletin, 31*, 1–9.

Deng CL (2010) Theory and application for generalized linear models. Thesis, Guangxi University, Nanning, Guangxi, China.

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z., Knowler, D. J., Leveque, C., Naiman, R. J., Prieur-Richard, A. H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews, 81*, 163–182.

Fan, J. T., Li, J. S., Xia, R., Hu, L. L., Wu, X. P., & Li, G. (2014). Assessing the impact of climate change on the habitat distribution of the giant panda in the Qinling Mountains of China. *Ecological Modelling, 274*, 12–20.

Froidevaux, J., Zellweger, F., Bollmann, K., Jones, G., & Obrist, M. K. (2016). From field surveys to LiDAR: shining a light on how bats respond to forest structure. *Remote Sensing of Environment, 175*, 242–250.

Fu, Y. Q., Zhang, Z. W., Chen, B. P., & Ling, Z. W. (2011). Winter habitat characteristics of red-winged laughing thrush at Laojunshan National Nature Reserve in China. *Chinese Journal of Zoology, 46*, 48–54.

Gaisler, J., Zukal, J., Rehak, Z., & Homolka, M. (1998). Habitat preference and flight activity of bats in a city. *Journal of Zoology, 244*, 439–445.

Gao, H., Tan, L. M., Liu, P., Liu, J. T., & Li, X. R. (2017). Estimation of arable land resources potential in hilly area of Taihang Mountain based on binary logistic regression model. *Chinese Journal of Eco-Agriculture, 25*, 490–497.

Gerald, K., & Markus, M. (2009). Species-specific barrier effects of a motorway on the habitat use of two threatened forest-living bat species. *Biological Conservation, 142*, 270–279.

Goiti, U., Garin, I., Almenar, D., Salsamendi, E., & Aihartza, J. (2015). Foraging by Mediterranean horseshoe bats (Rhinolophus euryale) in relation to prey distribution and edge habitat. *Journal of Mammalogy, 89*, 493–502.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley ISBN 978-047-172-214-4.

Hurvich, C. M., Shumway, R., & Tsai C. (1990). Improved estimators of Kullback-Leibler information for autoregressive model selection in small samples. *Biometrika, 77*, 709–719.

IUCN. (2016). The IUCN red list of threatened species. Version 2016-1. Retrieved from http://www.iucnredlist.org. Accessed 23 Oct 2016.

Ji, R., Gao, Z. X., Xie, B. Y., Li, D. M., & Zeng, X. C. (2007). Prediction on oviposition site selection Locusta migratoria manilensis by logistic model in coastal areas. *Acta Ecologica Sinica, 27*, 5029–5037.

Jiang GS (2007) Ecological effects of human disturbances on moose and roe deer and their adaptation mechanisms at multiple spatial scales. Dissertation, Northeast Forestry University, Haerbin, Heilongjiang, China.

Kalda, O., Kalda, R., & Liira, J. (2015). Multi-scale ecology of insectivorous bats in agricultural landscapes. *Agriculture, Ecosystems and Environment, 199*, 105–113.

Kelly, R. M., Kitzes, J., Wilson, H., & Merenlender, A. (2016). Habitat diversity promotes bat activity in a vineyard landscape. *Agriculture, Ecosystems and Environment, 223*, 175–181.

Li, Y. C., Meng, Y. H., Zhang, L. C., & Ye, Q. (2005). Analysis of environmental factors on geographical distribution of Chinese Chiroptera. *Acta Zoologica Sinica, 51*, 413–422.

Li, H., Gong, Z. N., Zhao, W. J., & Gong, H. L. (2012). Driving forces analysis of reservoir wetland evolution in Beijing based on logistic regression model. *Acta Geographica Sinica, 67*, 357–367.

Li, G. Q., Liu, C. C., Liu, Y. G., Yang, J., Zhang, X. S., & Guo, K. (2013). Advances in theoretical issues of species distribution models. *Acta Ecologica Sinica, 33*, 4827–4835.

Li, C., Yang, D. D., Zhang, Y. M., Song, Y. C., Li, P. F., & Jiang, Z. G. (2016). Seasonal variation in nocturnal bed-site selection by Milu (Elaphurus davidianus) in Hubei Shishou Milu National Nature Reserve, China. *Biodiversity Science, 24*, 1031–1038.

Liang, L., Di, L. P., Zhang, L. P., Deng, M. X., Qin, Z. H., Zhao, S. H., & Lin, H. (2015). Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. *Remote Sensing of Environment, 165*, 123–134.

Liang, L., Qin, Z. H., Zhao, S. H., Di, L. P., Zhang, C., Deng, M. X., Lin, H., Zhang, L. P., Wang, L. J., & Liu, Z. X. (2016). Estimating crop chlorophyll content with hyperspectral vegetation indices and the hybrid inversion method. *International Journal of Remote Sensing, 37*, 2923–2949.

Liang, L., Sun, Q., Luo, X., Wang, J. H., Zhang, L. P., Deng, M. X., Di, L. P., & Liu, Z. X. (2017). Long-term spatial and temporal variations of vegetative drought based on vegetation condition index in China. *Ecosphere, 8*(8), e01919.

Liu, W., Wang, Y. X., He, X. H., & Niu, H. X. (2011). Distribution and analysis of importance of underground habitats of cave-dwelling bats in the south of Taihang Mountain. *Acta Theriologica Sinica, 31*, 371–379.

Lumsden, L. F., Bennett, A. F., & Silins, J. E. (2002). Location of roosts of the lesser long-eared bat Nyctophilus geoffroyi, and Gould's wattled bat Chalinolobus gouldii, in a fragmented landscape in south-eastern Australia. *Biological Conservation, 106*(2), 237–249.

Mcconnell, S., O'Connell, T. J., & Leslie, D. M. (2008). Land cover associations of nesting territories of three sympatric Buteos in shortgrass prairie. *Wilson Journal of Ornithology, 120*(4), 708–716.

Miková, E., Varcholová, K., Boldogh, S., & Uhrin, M. (2013). Winter diet analysis in Rhinolophus euryale, (Chiroptera). *Central European Journal of Biology, 8*, 848–853.

Montechiesi, L., Cocconcelli, M., & Rubini, R. (2016). Artificial immune system via Euclidean distance minimization for anomaly detection in bearings. *Mechanical Systems and Signal Processing, 76-77*, 380–393.

Ouyang, Z. Y., Liu, J. G., Xiao, H., Tan, Y. C., & Zhang, H. M. (2011). An assessment of giant panda habitat in Wolong Nature Reserve. *Acta Ecologica Sinica, 21*, 1869–1874.

Qin, A. L., Liu, B., Guo, Q. S., Bussmannn, R. W., Ma, F. Q., Jian, Z. J., Xu, G. X., & Pei, S. X. (2017). Maxent modeling for predicting impacts of climate change on the potential distribution of Thuja sutchuenensis, Franch. An extremely endangered conifer from southwestern China. *Global Ecology and Conservation, 10*, 139–146.

Shu, X. J., Chen, Y. B., & Ren, Q. W. (2010). Model selection criteria for flood frequency analysis. *Journal of Hydraulic Engineering, 41*, 80–85.

Sukumal, N., Gale, G. A., & Savini, T. (2010). Sub-montane habitat selection by a lowland pheasant. *The Raffles Bulletin of Zoology, 58*, 391–401.

Uhrin, M., Benda, P., & Kaňuch, P. (2017). Different responses of attic-dwelling bat species to landscape naturalness. *Mammalian Biology, 82*, 48–56.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of Earth's ecosystems. *Science, 277*, 494–499.

Wang, Y. X. (2003). *A complete checklist of mammal species and subspecies in China, a taxonomic and geographic reference*. Beijing: China Forestry Publishing House ISBN 978-750-383-131-7.

Wang J (2009) Habitat selection and prediction of spatial distribution in greater horseshoe bats, Rhinolopinus ferrumquinum at multiple spatial scales. Dissertation, Northeast Normal University, Changchun, Jilin, China.

Wang F (2010) Phylogeny and biogeography of the greater horseshoe bat in Henan. Thesis, Henan Normal University, Xinxiang, Henan, China.

Wang, M. J., & Li, J. Q. (2008). Research on habitat restoration of giant panda after a grave disturbance of earthquake in Wanglang Nature Reserve, Sichuan Province. *Acta Ecologica Sinica, 28*, 5848–5855.

Wang, Y. S., Xie, B. Y., Wang, F. H., Xiao, Q. M., & Dai, L. Y. (2007). Application of ROC curve analysis in evaluating the performance of alien species' potential distribution models. *Biodiversity Science, 15*, 365–372.

Wu, Z. J., He, G. J., Huang, S. L., Wang, M. M., & Lin, J. T. (2017). Terrain effects assessment on remotely sensed fractional vegetation cover in hilly area of southern China. *Journal of Remote Sensing, 21*, 159–167.

Xie, L. F., Sun, K. P., Jiang, T. L., Liu, S., Lu, G. J., Jin, L. R., & Jiang, F. (2017). The effects of cultural drift on geographic variation in echolocation calls of the Chinese rufous horseshoe bat (Rhinolophus sinicus). *Ethology, 123*, 532–541.

Xu HL (2012) Sequence and organization of the complete mitochondrial genome of *Hipposideros armiger*, rhinolophus luctus, and *Petaurista alborufus*. Dissertation, Sichuan Agricultural University, Ya'an, Sichuan, China.

Xu, Z. L., Peng, H. H., & Peng, S. Z. (2015). The development and evaluation of species distribution models. *Acta Ecologica Sinica, 35*, 557–567.

You, J., Leng, Z. M., & Ding, J. J. (2013). *The development report of China contiguous destitute areas*. Beijing: Social Sciences Academic Press.

Yu, W. H., Fan, F. L., Jiang, H. S., Li, S. Y., Kong, X. K., & Wu, Y. (2015). Habitat selection of cave-bats in summer in Wuzhi Mountain, Hainan. *Guangdong Forestry Science and Technology, 31*, 74–79.

Zhang, X. N., Wang, S. W., Feng, J., Hu, X. L., Liu, Y., He, X. H., & Shi, L. M. (2008). Recognition method of wavelet packets for echo-locating calls of flying rhinolophus lepidus and rhinolophus rouxi. *Technical Acoustics, 27*, 44–48.

Zhang, D. D., Zhu, H. Q., Jiang, C. Y., Li, C., Zhang, X. D., Liu, C. F., Qu, D. J., & Guo, H. (2014). Research situation of impact factors on habitat selection of Siberian moschus moschiferus. *Journal of Economic Animal, 18*, 44–46.

Zhou, Z. M., Yang, Y. M., & Chen, B. Q. (2017). Estimating the Spartina alterniflora fractional vegetation cover using high spatial resolution remote sensing in a coastal wetland. *Acta Ecologica Sinica, 37*, 505–512.

Zribia, M., Hegarat-Mascle, S. L., Taconet, O., Ciarletti, V., & Vidal-Madjar, D. (2003). Derivation of wild vegetation cover density in semi-arid regions: ERS2/SAR evaluation. *International Journal of Remote Sensing, 24*, 1335–1352.