

Modelling the Spatial Distribution of Wildlife Animals Using Presence and Absence Data

H.-S. KWON¹, B.-J. KIM¹, G.-S. JANG²

¹ National Institute of Ecology
Seocheon, Chungnam 33657, South Korea

² Department of Life Sciences, Yeungnam University
Gyeongsan, Gyeongbuk 38541, South Korea
E-mail: sunside@ynu.ac.kr.

Статья поступила 24.12.15

Принята к печати 30.12.15

ABSTRACT

This study was conducted to analyze the habitat preference of six major mammals for various environmental factors based on 100 random points within a mountain area in South Korea. In-situ presence and absence data for the mammals were surveyed and collected, and twelve explanatory variables related to topography, water, greenness, and anthropogenic influence were applied to create a habitat distribution model. The best combination of variables was determined using Moran's I coefficients and Akaike criteria information, and applied to estimate the habitat preference for each species using GRASP v.3.0. The predictive map showed that wildlife animals in this study were mainly categorized into two groups: Group I (Korean squirrel, *Sciurus vulgaris*, mole, *Talpa micrura* and water deer, *Hydropotes inermis*), showed equal preference for all mountainous areas; Group II (weasel, *Mustela sibirica*, leopard cat, *Felis bengalensis* and raccoon dog, *Nyctereutes procyonoides*) showed different preferences in a mountain.

Key words: Species distribution model, General additive model, GRASP, Habitat preference.

Interest in species distribution models (SDMs) of plants and animals has been growing dramatically and they have become increasingly important tools to address various issues in ecology, biogeography, evolution and climate change since the early 1990s [Guisan, Thuiller, 2005]. More recently, prediction of species distributions through field-surveying data has long been recognized as a significant component of conservation planning [Austin, 2002; Elith, Burgman, 2003], and a wide variety of statistical and machine-learning methods have

been introduced to obtain the best predictions, often in conjunction with geographic information systems (GIS) and remote-sensing technology [Franklin et al., 2000].

Recently, various forms of regression analyses have been applied to predicting species distribution. Generalized linear models (GLM) [McCullagh, Nelder, 1989; Friston et al., 1995], generalized additive models (GAM) [Hastie, Tibshirani, 1990] and Maximum entropy (Maxent) [Phillips et al., 2006] are becoming increasingly popular for prediction of species dis-

tribution [Scott et al., 2002]. GLMs and GAMs are commonly used for modeling with presence/absence datasets, although they have been applied to presence-only situations by taking a random sample of pixels from the study area, known as “background pixels” or “pseudo-absences” [Ferrier et al., 2002]. GAM, an extended GLM, is well known to efficiently explain the non-linear relationship between dependent and independent variables [Elith et al., 2006; Franklin, 2010]. GAMs are non-parametric extension of GLMs, which are themselves a generalization of classical least square regression (LSR). While GLMs extend the application of classical regression into other statistical distributions (binomial, Poisson, gamma, negative binomial), GAMs estimate response curves with a non-parametric smoothing function instead of parametric terms (e. g. $ax + bx^2$) [Lehmann et al., 2002]. This allows exploration of shapes of species response curves to environmental gradients, and enables fitting of statistical models in better agreement with ecological theory [Austin, 2002]. Maxent is derived from a branch of statistics referred to as machine learning [Yackulic et al., 2012], and is a general-purpose method for making predictions or inferences from incomplete information [Phillips et al., 2006]. Although the Maxent can be applicable to the model using presence/absence data, it was well known to get much higher confidence in the presence-only model [Kim et al., 2012; Elith et al., 2006].

Generally presence/absence data or presence-only data are used as a binary response variable for modeling the species distribution. However, the process of getting presence/absence data is quite different from the one for presence-only data. The presence/absence data for target species are surveyed and determined at randomly located points in a field, while presence-only data for an interested species are obtained from a geographically registered relational database of incidental sightings of the species [Gormley et al., 2011]. Although several investigations of habitat modeling have adopted regression analyses using currently available presence-only data in conjunction with pseudo-absence data created at random by computer simulations, presence-only data are known problematic for several analyses because

owing to data limitations, analyses of species distribution are limited to areas of small spatial extent [Ferrier, Guisan, 2006]. Thus, there is a need for spatially statistical models with in-situ presence/absence data to enable the proper estimation and interpretation of the occurrence of specific species. Therefore, this study was conducted to model the spatial distribution of wildlife mammals using in-situ presence/absence data surveyed at 100 random sites in Dangjin, South Korea.

MATERIALS AND METHODS

Study area. Dangjin is a small city located on the west coast of South Korea (Fig. 1). The population of Dangjin has been growing continuously since 1980, and it officially exceeded 150,000 in 2014 [Dangjin city..., 2014]. Because of the rapid pace of urbanization, broad-scale habitat destruction is widespread across the region, and much of the original forest exists in fragmented patches.

Rapid urbanization has been driving degradation of forest patches, and 10 km² of forests were removed 10 years ago, which was coupled with an approximately 10 km² increase in building sites. During the same period from 2006 to 2012, 52.9 km² of roads were newly constructed and paved. Accordingly, many forest patches have undergone fragmentation and decreases in quality. Thus, urgent investigation is needed to preserve the natural environment in this city, and information regarding wildlife animals obtained through such research is expected to facilitate creation of policies protecting natural resources in Dangjin.

In this study, 100 points were randomly generated in mountain areas to survey the presence or absence of wildlife mammals. Although many previous researchers have investigated the effects of sample size on model accuracy, most have not explicitly manipulated sample size by species [Brotons et al. 2004]. Hernandez et al. [2006] generated models for 18 California taxa with varying degrees of habitat specialization using four modeling methods and a variety of sample sizes characteristic of rare species ($n = 5, 10, 25, 50, 75, \text{ or } 100$) to further explore variations among sample size and model performance. Wisz et al. [2008] also

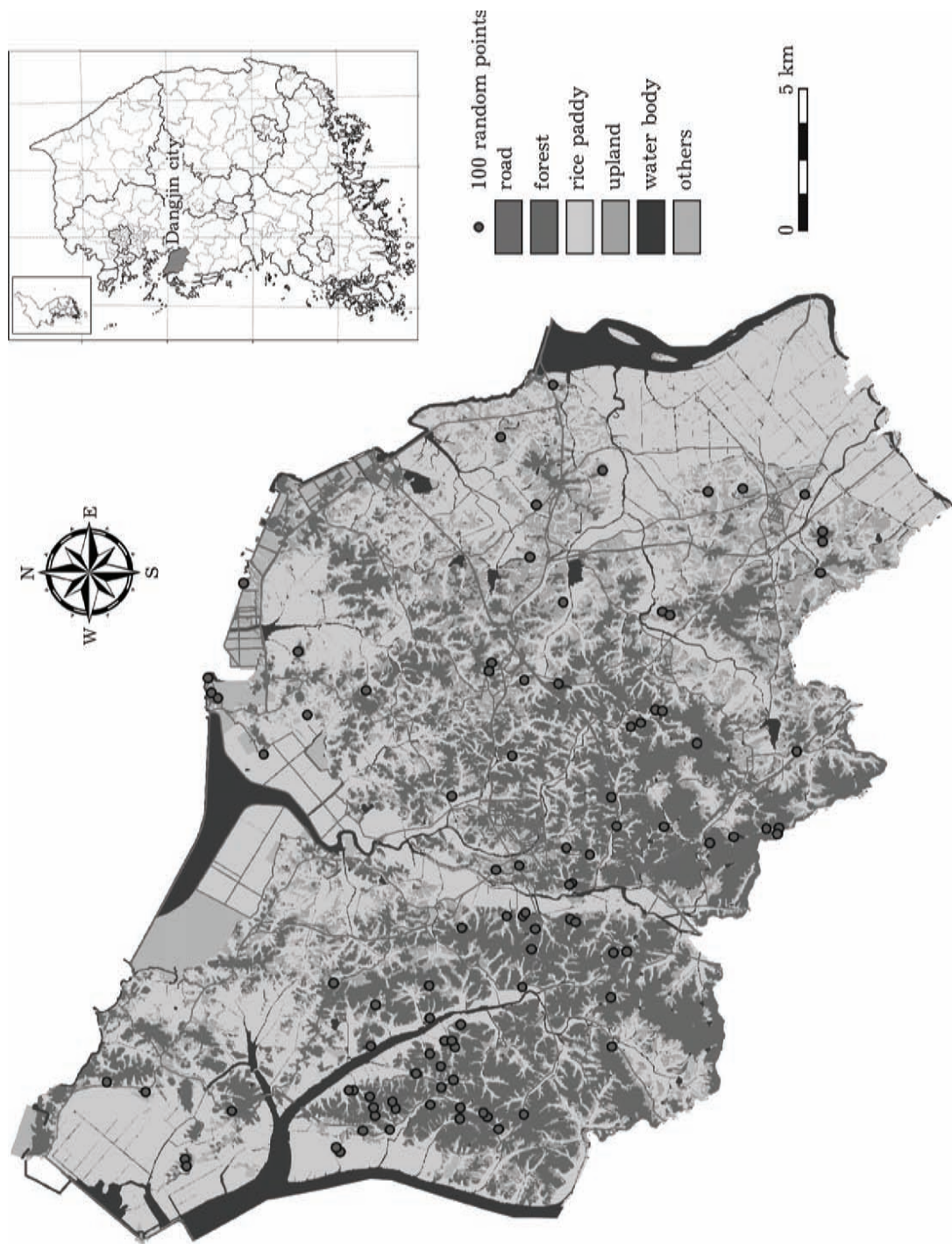


Fig. 1. The study region in Dangjin City area

evaluated predictions from 12 algorithms for 46 species in three sample pools (100, 30 and 10 records) and found that model accuracy decreased and variability increased across species and between models with decreasing sample size. The aforementioned studies would guarantee the number of sampling sites set in this study were sufficient to explain the species distribution, although the maximum accuracy potential and the sample size at which the asymptote is reached will depend on the study area and species.

The create random points tool in ArcMap v.9.3 (ESRI, CA, USA) was used to obtain 100 random points generated within the forest polygons in Dangjin to survey the occurrence of wildlife mammals. A specified number of random point features can be generated in ArcMap, in which the create random points tool can be used to generate random points within the area defined by the constraining polygon.

Data collection for presence/absence data of target species. Field surveys were conducted at 100 random points two times, once in May and again in September of 2008, each site was traced using GeoXT GPS (Trimble, CA, USA) embedded with a shape file containing the location of 100 points. Each survey was conducted in the morning (6:00 am to 12:00 noon) and late afternoon (4:00 pm to 7:00 pm) for 20 days. Most surveys of wildlife animals are best undertaken at dawn or dusk and not during summer or in hot weather [Eldridge, Pearson, 2008]. A circular quadrat was established within a 25-meter radius of each point, because using circular quadrats instead of square or rectangle ones minimizes the edge effect of a quadrat [Grimbacher, Catterall, 2007]. The survey was carried out by recording all of the mammals that were observed

and their traces (e. g., stools and footprints) left in each site. The occurrence data from the two surveys were then recorded and listed in a Table 1. Overall, nine species of wildlife mammals were surveyed at the 100 sites; however, *Tamias sibiricus*, *Micromys minutus* and *Lepus sinensis* were excluded from further analyses because their frequency was too low to create a proper model.

Information regarding the occurrence of mammals at each site was converted to binary data, with 1 indicating their presence and 0 their absence. An observed absence is usually employed when the survey method fails to detect a species that actually resides at or uses the location, because it is generally impossible to confirm a species' absence [MacKenzie, 2005]. The presence/absence data at the 100 points were used as response variables for the species distribution model. The six mammals surveyed in this study were representative species living in South Korea, and the models estimating their occurrence would be applicable to other sites in South Korea.

Variables explaining the habitat preference of mammals. Most habitat models are generally based on various hypotheses describing how environmental factors control the distribution of species and communities. In this study, several categories of environmental factors were considered from the perspective of a habitat adopted by Grillet et al. [2010] for habitat modeling. Food, cover, water and places to live were the four main habitat components considered for the model.

Forests consist of food, cover and areas for wildlife survival [McCollin, 1998]. Farmland and vegetation strips near rivers also contain these habitat components [Darveau et al., 1995]. These landscape components are recognized as im-

T a b l e 1

Terrestrial vertebrates and their frequency surveyed at 100 sites in this study

Name of species			Name of species		
English name	Scientific name	Frequen- cy	English name	Scientific name	Frequen- cy
Water Deer	<i>Hydropotes inermis</i>	54	Hare	<i>Lepus sinensis coreanus</i>	7
Raccoon dog	<i>Nyctereutes procyonoides</i>	24	Leopard Cat	<i>Felis bengalensis euphilura</i>	15
Asiatic chipmunk	<i>Tamias sibiricus</i>	1	Weasel	<i>Mustela sibirica coreana</i>	12
Mole	<i>Talpa micrura coreana</i>	58	Korean squirrel	<i>Sciurus vulgaris coreae</i>	63
Harvest Mouse	<i>Micromys minutus</i>	4	Total number	238	

portant factors for the survival of wildlife animals, and species preferring a specific land cover could be susceptible to reductions and fragmentation of land cover, which result in reductions in habitat [Andrén, 1994; Fahrig, 2002].

R. D. Mace et al. [1999] attempted to estimate leaf greenness using the reflectance derived from a Landsat Thematic satellite image, while A. Guisan and N. E. Zimmermann [2000] obtained an environmental variable from a satellite image to evaluate the quality of forests. The normalized difference vegetation index (NDVI) derived from a Landsat ETM (Enhanced Thematic Mapper) taken in May 5, 2008 was used to discriminate the different levels of greenness in forests. The NDVI values (-1 to 1) were converted to 8-bit unsigned integer format ranging from 0 to 255.

Many studies related to habitat modeling have adopted topographic factors as indicators for evaluation of wildlife habitat [Guisan, Zimmermann, 2000]. Elevation, slope and aspect have commonly been used for modeling because different levels of solar energy could be radiated according to these topographic conditions, resulting in generation of different habitat conditions [Dubayah, Rich, 1995]. In this study, environmental variables related to topographic factors were obtained from a digital topographic map drawn on a scale of 1 to 5,000 (National Geographic Information Institute, Suwon, South Korea). A digital elevation model created from the digital topographic map was then converted into slope, relief, curvature, northness and wetness.

Among human activities that cause habitat loss, urban development produces some of the greatest local extinction rates and frequently eliminates the majority of native species. Negative factors such as buildings and roads are also needed to accurately estimate the wildlife habitat. Tratalos et al. [2007] adopted the concept of urban form as an index to evaluate biodiversity, which was quantified by building density per hectare. Linear features in landscapes such as roads can also affect habitat fragmentation, ultimately influencing the distribution pattern of species in a habitat [Vos et al., 1998; Cushman, 2006]. The ratio of buildings or roads per hectare was calculated using the Korean Land Information System (KLIS) managed by

the Korean Ministry of Land, Infrastructure and Transport and applied to determine the influence of buildings and roads on wildlife habitats.

Water has long been recognized as one of the most important resources in a habitat for all organisms, including humans [Lawton, Woodroffe, 1991; Gibbons, 2003]. Additionally, some species show extremely preference for humidity, while others do not depend greatly on humidity; therefore, the different humidity associated with nearby water bodies can cause an environmental gradient [Cornwell, Ackerly, 2009] and disturb natural ecosystems [Patz et al., 2000]. In this study, distance from water bodies was used to explain preferences of each species for various distances to water bodies.

Spatial autocorrelation by Moran's I. Measurement of the magnitude and direction of spatial autocorrelation is needed for exploratory analysis of spatial data. Specifically, this has been defined as "the property of random variables taking values, at pairs of locations being a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations" [Legendre, 1993]. Moran's I coefficient, which is one of the most commonly used descriptors of spatial autocorrelation, is an index measuring spatial autocorrelation based on both feature locations and feature values. Given a set of features and an associated attributes, it evaluates whether the pattern expressed is clustered, dispersed, or random. SAM (Spatial Analysis in Macroecology) version 4.0 is useful to describe the spatial pattern in the distribution of species by Moran's I coefficient [Rangel et al., 2010]. This software was used in this study to measure the magnitude of spatial autocorrelation at 100 points in each environmental variable.

Species Distribution Model. GAM was employed in this study to estimate the spatial distribution of 6 major mammals in Dangjin, South Korea. Alternative multivariate models were generated by stepwise GAM in GRASP v. 3.0 [Lehmann et al., 2002; <http://www.inside-r.org/packages/cran/grasp>], and the best model was selected using the Akaike Information Criterion, AIC [Burnham, Anderson, 2001;

Thuiller, 2003]. GRASP (Generalized Regression Analysis and Spatial Prediction) is plugged in R or S-Plus statistical packages (<http://www.inside-r.org/packages/cran/grasp>) that offer an automated way of making spatial predictions from point surveys using Generalized Additive Models. AIC is a criterion that H. Akaike [1974] developed for model selection and evaluation based on the S. Kullback and R. A. Leibler [1951] information of the true model with respect to the fitted model. The AIC value indicates the proper model minimizing the information loss estimated. In this study, backward selection, one of the stepwise selection methods, was used to select the proper variables best explaining model. This information was estimated to minimize information loss by referring to the AIC value.

The accuracy of each model was calculated by the AUC (area under the cover) derived from a ROC (receiver operating characteristic) curve. The accuracy using the AUC is fixed independently of the threshold set for the model, while the accuracy of other models is changeable depending on its threshold [Thuiller, 2003]. To determine the probability from habitat modeling in this study, it was necessary to convert the threshold distribution pro-

bability to the estimation model based presence/absence data. To create binary maps from probability maps, 0.5 is often used in logistic regression and related modeling [Manel et al., 1999; Franklin, 2010]. Generally, a probability < 0.5 represents being absent, while a higher probability > 0.5 represents being present. Some other methods referred to the standard deviation from the distribution probability. This study constructed a binary map using a threshold, which maximized the probability by correctly estimating presence and absence data [Franklin, 2010].

RESULTS

Spatial autocorrelation of each environmental variable. Moran's I coefficients showed a lower spatial autocorrelation at pairs of points for all environmental variables except elevation (Table 2, 3). Most variables showed Moran's I coefficients with an absolute value not higher than 0.2. These findings indicated that the properties of each variable at pairs of locations were expected for randomly associated pairs of observations. These data suggested that adjacent habitat points did not share similar values, so it was assumed that adjacent points

T a b l e 2
Explanatory variables used for the habitat prediction model

Acronym of variables	Unit	Description of each environmental variables	Data type
Elev	m	Elevation: using digital elevation model	Continuous
Slope	degree	Slope degree: arctan (rise / run)	Ibid
Relief	–	Standard deviation of elevation (5x5 windows)	»
Curv	–	Curvature: for each cell, a fourth-order polynomial of the form	»
Asp_1	degree	Aspect: ranging from 0 to 360 degree	»
Asp_2	degree	Aspect (ranging from 0 to 180 degree)	»
North	–	Northness: $\cos \{(\text{aspect} \cdot \pi)/180\}$, $360^\circ = 1$ and $180^\circ = 0$	»
Wet	–	Wetness: $\ln\{(\text{flow accumulation} + 1)/(\text{slope} + 1)\}$	»
NDVI	–	Normalized Difference Vegetation Index using Landsat ETM+ taken in May 5, 2008	»
Bldg.d	%	Building density: building areas per ha	»
D.farm	m	Distance from farm lands	»
D.bldg	m	Distance from buildings	»
D.road	m	Distance from roads extracting from digitalized topographic map	»
D.water	m	Distance from water bodies, which was derived from digitalized topographic map	»

T a b l e 3

Moran's I coefficients for each variable

	Count	Moran's I	P	(max)	I/I(max)
Asp_1	776	-0.004	0.93	0.415	-0.01
Asp_2	776	-0.046	0.302	0.483	-0.095
Curv	776	-0.049	0.307	0.382	-0.127
North	776	0.019	0.704	0.589	0.032
Relief	776	0.013	0.839	0.717	0.018
Slope	776	0.075	0.136	0.757	0.099
Elev	776	0.372	0.005	0.727	0.512
NDVI	776	0.035	0.487	0.543	0.065
Wet	776	0.058	0.216	0.525	0.11
Bldg. d	776	0.082	0.095	0.58	0.141
D. bldg	776	0.177	0.01	1.012	0.175
D. farm	776	0.036	0.472	0.748	0.048
D. road	776	0.003	0.965	0.551	0.005
D. water	776	-0.069	0.08	0.345	-0.2

randomly sampled were independent based on the spatial autocorrelation test [Koenig, 1999].

Selecting variables for the best model and model validation. Variables were checked by backward stepwise selection using the AIC at each step (Table 4). Variables were removed step by step by determining which model with remaining variables could be best explained by the lower AIC, or lower unexpected deviance. A smaller AIC indicates a better model; there-

fore, there is greater deviance explained per number of explanatory variables (parameters estimated) [Franklin, 2010].

The best combination of variables was determined and applied to estimate the habitat preference for each species using GRASP v.3.0 (Table 5). Variables selected for each species showed higher AUC values, indicating that each model explained the distribution of each species well. *F. bengalensis* had the highest

T a b l e 4

Variables removed by the backward stepwise selection method and AIC for each model

Steps	<i>S. vulgaris</i>		<i>F. bengalensis</i>		<i>T. micrura</i>		<i>N. procyonoides</i>		<i>H. inermis</i>		<i>M. sibirica</i>	
	Var	AIC	Var	AIC	Var	AIC	Var	AIC	Var	AIC	Var	AIC
1		134		88		142		134		162		92
2	Asp_2	129	Relief	82	D.road	137	Asp_2	129	Asp_1	157	D.bldg	86
3	Curv	125	Wet	77	Asp_1	132	Slope	124	Wet	151	Curv	82
4	D. water	122	Slope	73	North	128	North	119	Asp_2	147	Wet	77
5	Bldg. d	117	Asp_2	68	Curv	124	Wet	115	North	142	D.farm	73
6	D. bldg	114	North	65	D.farm	120	Asp_1	112	D. farm	138	Asp_1	70
7	Elev	112	Curv	62	Elev	116	Curv	108	Slope	134	Elev	68
8	North	109	D. road	59	Bldg. d	113	Elev	106	Bldg. d	131	Bldg. d	66
9	Slope	106	NDVI	58	Wet	113	D. road	105	NDVI	127	Relief	66
10	NDVI	104					Relief	104	Elev	124	D. road	64
11									D. road	121	North	63
12									D. water	120		

N o t e. AIC value was rounded off below the decimal poin.

Summary of the best combination of variables for GAM models to predict presence and absence of each species

Species	Selected variables (order)	AUC	cvAUC
<i>S. vulgaris</i>	D. road, D. farm, Wet, Relief, Asp_1	0.890	0.731
<i>F. bengalensis</i>	D. water, D. farm, D. bldg, Bldg. d, Elev, Asp_1	0.996	0.839
<i>T. micrura</i>	Slope, D. water, Relief, NDVI, Asp_2, D. bldg	0.870	0.745
<i>N. procyonoides</i>	D. water, D. farm, D. bldg, Bldg. d, NDVI	0.889	0.674
<i>H. inermis</i>	D. bldg, Relief, Curv	0.793	0.631
<i>M. sibirica</i>	NDVI, Slope, D.water, Asp_2	0.945	0.792

N o t e. AUC: area under the curve, cvAUC: cross-validated AUC.

cvAUC (Cross-Validated Area Under the ROC Curve), indicating that the variables selected with higher cvAUC values had greater significance and explained the habitat preference for the species well [Thuiller, 2003].

Predictive map. The predictive map showed that the wildlife animals in this study were primarily categorized into two groups. Group I

(*S. vulgaris*, *T. micrura* and *H. inermis*) showed higher preference for all locations on the mountain while Group II (*M. sibirica*, *F. bengalensis* and *N. procyonoides*) preferred specific areas on the mountain (Fig. 2). These findings indicate that regions with higher probability were distributed in almost all of the mountainous areas for Group I, while regions with higher

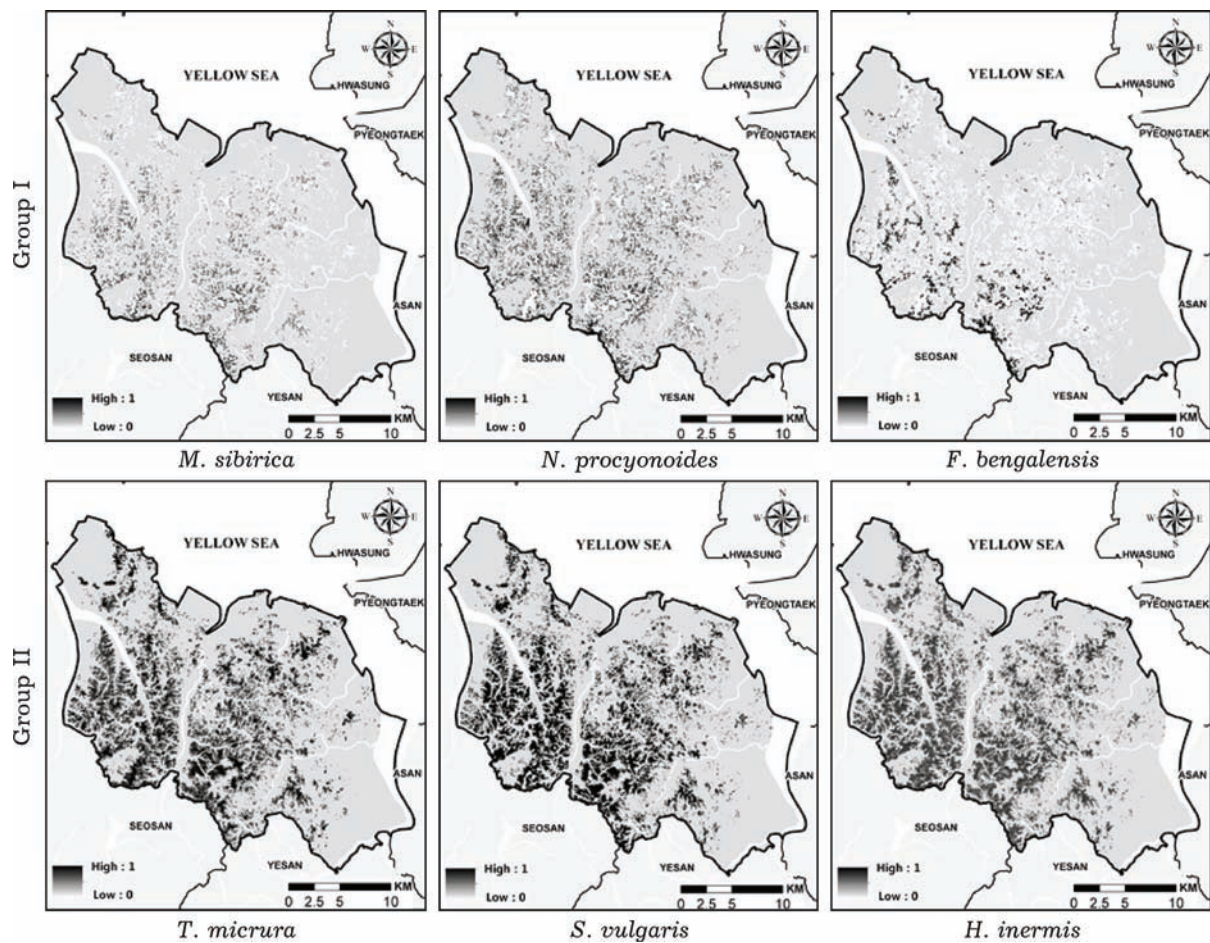


Fig. 2. Maps of occurrence probabilities predicted from generalized additive models (GAMs) for each species in Dangjin city, South Korea. Dark squares (30 × 30 m) represent areas of higher predicted occurrence

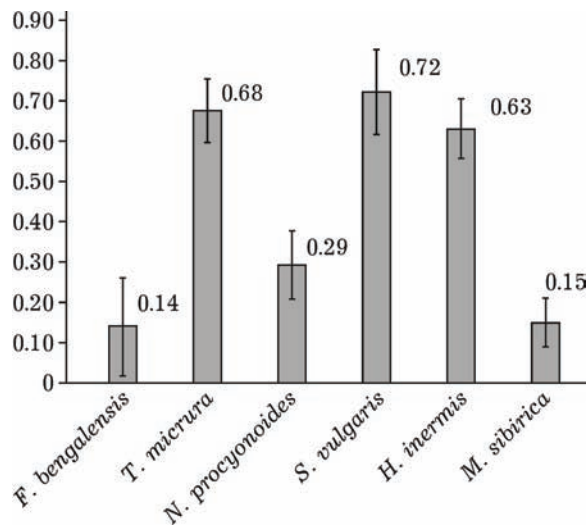


Fig. 3. Occurrence probability for 6 animals in average within forest areas

probability for Group II were sporadically distributed throughout mountains. Furthermore, regions with a higher occurrence probability for species included in Group I were likely to be coincident with forest boundaries in Dangjin, whereas regions for Group II were likely to be partially and independently distributed in mountains. These findings are reinforced by the data presented in Fig. 3, which indicate that the occurrence probability for Group I was relatively higher (>0.6), while the probability for Group II did not exceed 0.3. Species in Group I tend to prefer to live anywhere in mountainous areas, regardless of environmental conditions.

Partial probability response curves of generalized additive model (GAM) in forest area.

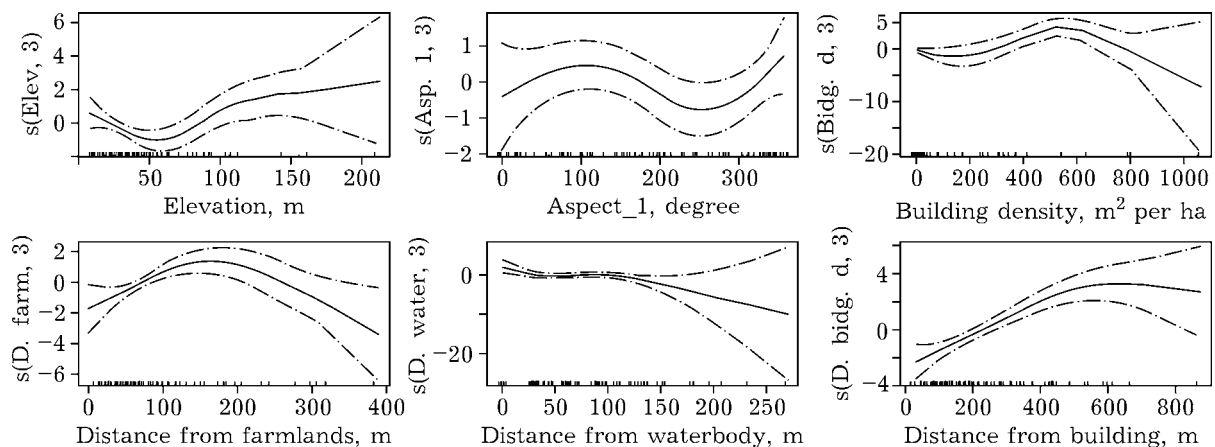


Fig. 4. Partial probability response curves for *F. bengalensis*

According to the partial probability response curves, *F. bengalensis* simultaneously preferred lower elevation (<50 m) and slightly higher elevation areas on the mountain (Fig. 4, a). These findings can be explained by the preference of *F. bengalensis* for the mountainous area as cover resources, while hunting prey in agricultural areas located at lower elevation. Moreover, the eastern slopes of a mountain are the preferred sites for this species (see Fig. 4, b). *F. bengalensis* did not prefer to be close to buildings (see Fig. 4, f), although this species was not sensitive to higher densities of buildings (around $500 \text{ m}^2/\text{ha}$; see Fig. 4, c). *F. bengalensis* appeared to depend on farmland (see Fig. 4, d) and bodies of water (see Fig. 4, e). As shown in fig. 4, c, d, this species was often found occurring at places close to farmlands (i. e. <200) and bodies of water (i. e. <150).

N. procyonoides had very different preferences for several environmental conditions. Specifically, this species did not appear to be sensitive to living close to buildings (Fig. 5, b), and it showed a higher frequency in areas with higher building density (see Fig. 5, a). Farmland (see Fig. 5, c) and bodies of water (see Fig. 5, d) were important and critical factors describing the preference of *N. procyonoides* for its habitat. Specifically, this species preferred areas close to farmland (i. e., <200) and bodies of water (i. e., <200). Moreover, the probability of its occurrence increased as the NDVI value increased (e. g., in forest and grassland).

H. inermis preferred fluctuant (Fig. 6, a) and concave (see Fig. 6, b) topography for its liv-

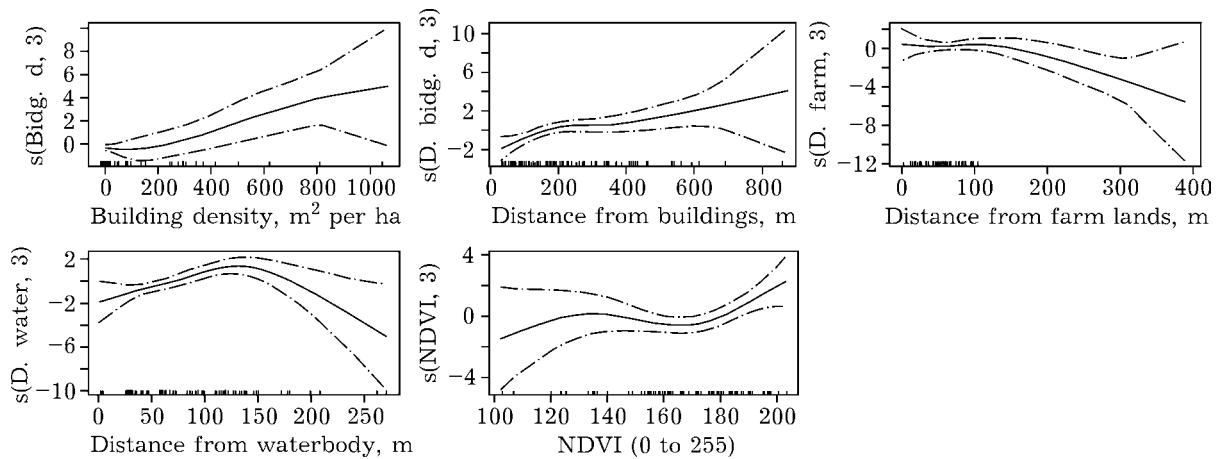


Fig. 5. Partial probability response curves for *N. procyonoides*

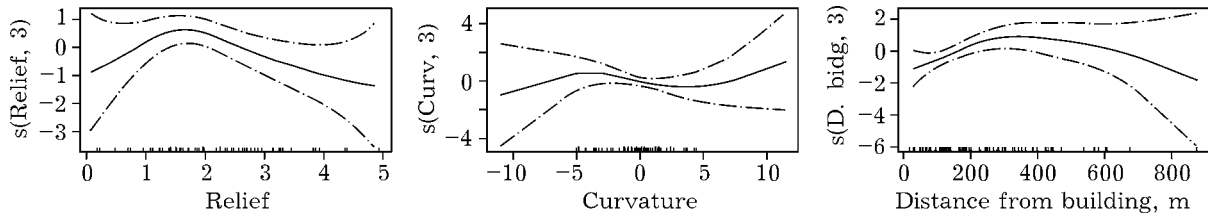


Fig. 6. Partial probability response curves for *H. inermis*

ing, and was often found at places not far from buildings and other artificial land uses (i. e., between 200 and 600 meters from buildings; see Fig. 6, c).

Slope clearly explained the preference of *M. sibirica* for its habitat, with its most often being present in areas with a gradient of around 20° (Fig. 7, a). NDVI also explained the preference of *M. sibirica* for its habitat, with a positive response to greenness (see Fig. 7, d).

M. sibirica was also shown to stay within approximately 200-m of water (see Fig. 7, c), although it was found everywhere, indicating that it was not dependent on this aspect (see Fig. 7, b).

Relief explained the occurrence probability of *S. vulgaris*. Specifically, it occurred at relief values of 1 to 3, indicating a fluctuant topography in the middle of the mountain area (Fig. 8, b). This species also preferred moun-

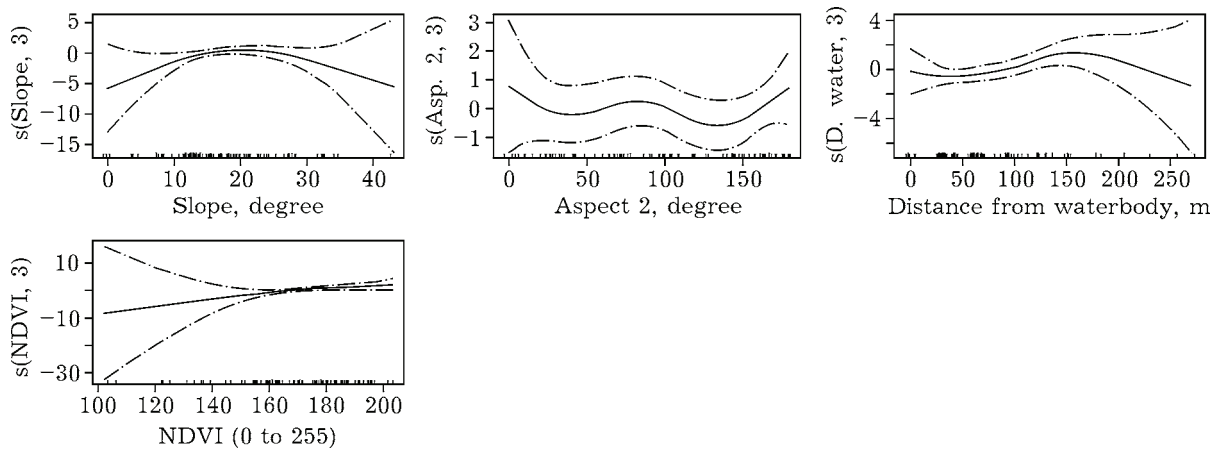


Fig. 7. Partial probability response curves for *M. sibirica*

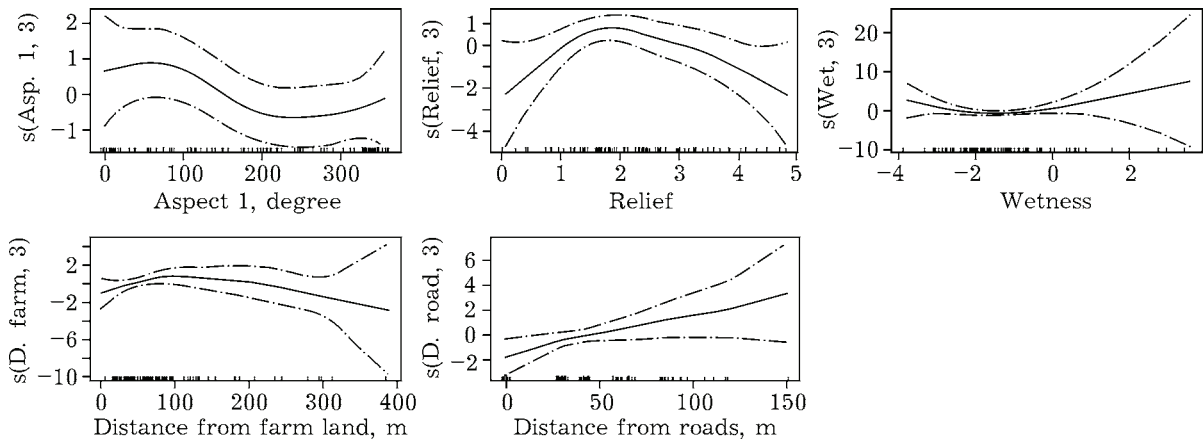


Fig. 8. Partial probability response curves for *S. vulgaris*

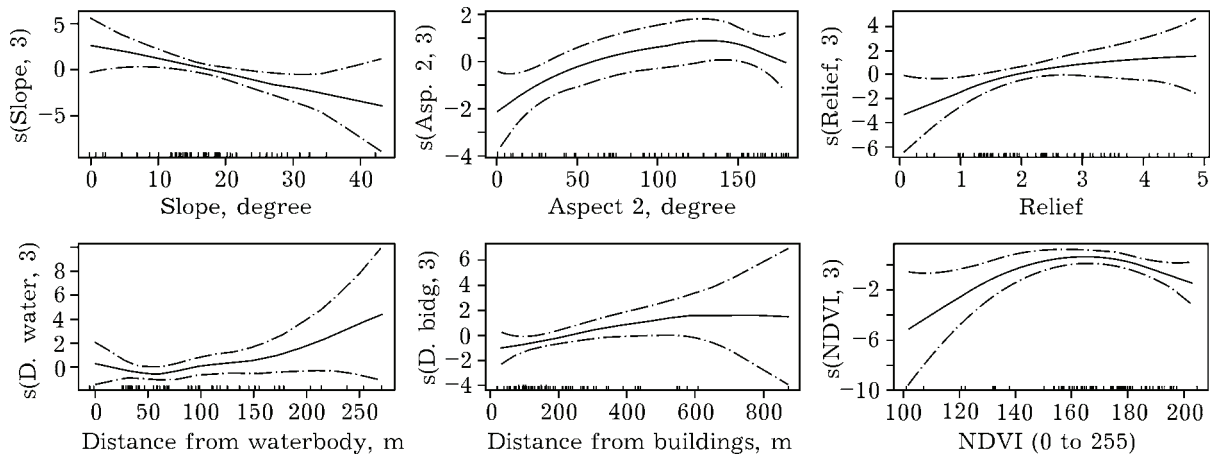


Fig. 9. Partial probability response curves for *T. micrura*

tain areas close to farmland (i.e., within about 200 meters of farmland) (see Fig. 8, d). Roads did not significantly influence the occurrence of *S. vulgaris*, although it was not found close to roads and it showed slightly higher preference for areas far from roads (see Fig. 8, e). As shown in Fig. 8, a, *S. vulgaris* appeared everywhere without any dependence on aspect, but it showed a slightly higher occurrence on the northeastern side of the mountain. Figure 8, c also explains the independence of *S. vulgaris* on wetness. Rugs, gathered between -3 and 0 on the x -axis in Fig. 8, c, were indicating the normal wetness indices appeared in a mountain area.

Slope was shown to be a geographical factor leading to a negative relationship with the occurrence of *T. micrura* (Fig. 9, a). This organism often appeared on gently sloped hills

of 10 to 30° , while it was not commonly found on slopes $> 40^\circ$. *T. micrura* occurred at reliefs of 1 to 3 , indicating the fluctuant area on a mountain (see Fig. 9, c). Moreover, it was found to prefer areas with higher NDVI values (> 150), indicating that this species preferred areas with sufficient greenness (see Fig. 9, f). As shown in Fig. 9, d, e, *T. micrura* was not highly sensitive to wetness and buildings.

DISCUSSION

The model using presence and absence data provided more accurate results than the one using presence-only data to explain the occurrence probability of wildlife animals. This was likely because of the higher uncertainty associated with the presence-only data for explaining the avoidance of a specific environmental

factor by a targeted animal without absence data obtained in the field.

Variables that remained after being removed by the backward stepwise selection were mainly associated with habitat components (i. e., “distance from bodies of water”, “NDVI” and “distance from farmland”) and man-made disturbances (i. e., “distance from buildings” and “building density”). These findings indicate that wildlife animals responded positively to environmental factors related to habitat components, while animals avoided disturbances caused by humans. Although several animals were not sensitive to the disturbance, most tended to respond to and/or avoid such disturbances.

S. vulgaris is known to be a significant predator of conifer seeds and disperser of oaks, indicating that this species is a resident in dominant coniferous forests and oak forests. The dominance of these plant species could be a cause of the widespread distribution of *S. vulgaris* on mountainous area. *T. micrura* is also a general species that appeared in all locations, regardless of any preference in space. Soil is deeper in areas with lower slope because the soil eroded from the upper slope positions (convex) is translocated downslope [De Alba et al., 2004; Papiernik et al., 2009]. These findings were similar to the response to greenness estimated based on the NDVI value in Fig. 9, *f*, in which the occurrence of *T. micrura* showed a plateau at around 160 in the 8-bit NDVI value that reflected the edge space of a mountain. These areas were found to include arable fields, deciduous woodlands, and permanent pastures, which all have soils deep enough to allow tunneling.

Fluctuant (see Fig. 6, *a*) and concave (see Fig. 6, *b*) topography are representative of the edge areas of a mountain, which would be favored by edge species such as *H. inermis* [Rhim, Lee, 2007]. *H. inermis* might require fluctuant and concave topography as a cover resource to hide from predators. Although *H. inermis* is known to be negatively affected by the presence of humans [Zhang et al., 2006], the species selected habitats near residential areas (see Fig. 6, *c*), which could be an alternative strategy for its survival. Indeed, predation of *H. inermis* calves by *F. bengalensis* and *N. procyonoides* have driven female

H. inermis to calve closer to roads and humans, which these predators are more apt to avoid [William et al., 2004].

There were no clear data describing the preference of *M. sibirica* for its selected habitat (see Fig. 7). Indeed, the results only indicated that *M. sibirica* preferred places not far from water (see Fig. 7, *c*) with a slightly higher slope (see Fig. 7, *a*) and sufficient greenness (see Fig. 7, *d*). These areas slightly overlapped with the six types of habitats (plain hedge, hedge with ditch, fence, drainage ditch, wood edge and riverbank) that D. W. Macdonald et al. [2004] reported as range compositions for *M. sibirica*. These findings indicate that the species would not enjoy the inner space in a mountain, which indicates they preferred out-of-mountain hills with a gentle slope and vegetation. Such areas would be optimal for these organisms because they enable the species to protect itself while enabling easy access to lowlands for hunting. After correction of the aforementioned estimations, the preferences derived from fig. 7 were coincident with its home range and habitat as described by H. Sakaki and Y. Ono [1994].

The preference of *F. bengalensis* for two extremes in elevation showed the potential for multiple habitats for its survival. Specifically, lower elevations were the best place for it to obtain prey, while higher elevations were needed to provide cover. This species appears to prefer the cover at higher elevation due to its sensitivity to anthropogenic pressure and competitors. This argument can be supported by Fig. 4, *f*, which shows that *F. bengalensis* did not prefer to be close to buildings. This strategy is quite different from that of *N. procyonoides*, which responded to building density.

N. procyonoides and *F. bengalensis* shared their habitats in places close to water bodies (see Fig. 4, *e*; 5, *d*) with some building density (see Fig. 4, *c*; 5, *a*). However, *N. procyonoides* was no more sensitive to the distance from buildings than *F. bengalensis*. Additionally, *F. bengalensis* was hesitant to get close to farms (see Fig. 4, *d*), while *N. procyonoides* entered farmland without any hindrance (see Fig. 5, *c*). Overall, these results indicate that *F. bengalensis* was much more cautious and sensitive to anthropogenic land uses than *N. procyonoides*.

CONCLUSION

The predictive map, showed that wildlife animals in this study, were mainly categorized into two groups according to their preferring environmental conditions. One group (Korean squirrel, *Sciurus vulgaris*, mole, *Talpa micrura* and water deer, *Hydropotes inermis*) was shown to prefer mountainous areas, while the other group (weasel, *Mustela sibirica*, leopard cat, *Felis bengalensis* and raccoon dog, *Nyctereutes procyonoides*) showed different preferences within a mountain. Different preferences in a mountain meant that each species might be dependent on different environmental conditions for each survival.

Variables used to explain the model for each species were primarily related to habitat components (distance from water bodies, NDVI and distance from farmland) and man-made disturbances (distance from buildings and building density). These findings indicate that wildlife animals responded positively to environmental factors related to the habitat components, while they avoided areas disturbed by humans, even though some were not sensitive to this disturbance. Especially, herbivores or mammals of small size like *S. vulgaris*, *T. micrura* and *H. inermis* didn't show any biased preference within a mountain area. Instead, they showed the higher preference for their habitat. On the other hand, mammals of relatively large size like *M. sibirica*, *F. bengalensis* and *N. procyonoides* had quite different strategies for their habitats. First of all, these large predators had the wider home range in a mountain so that they could get the flexibility to take their food within a mountain according to the presence of their food and/or cover resources. Among the predators, some species like *F. bengalensis* had quite biased preference on their habitat within a mountain rather than others like *M. sibirica* and *N. procyonoides* had. *F. bengalensis* seemed to prefer rather narrower range in mountainous area than others had because of their sensitivity to anthropogenic pressures. Hence, knowledge on the habitat preference of wildlife mammals in a mountain area nearby a developing city is important to protect themselves and their preferable habitats.

REFERENCES

- Akaike H. A new look at the statistical model identification // IEEE T Automat. Contr. 1974. Vol. 19. P. 716–723.
- Andrén H. Effect of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review // Oikos. 1994. Vol. 71. P. 355–366.
- Austin M. P. Spatial prediction of species distribution: an interface between ecological theory and statistical modeling // Ecol. Model. 2002. Vol. 157. P. 101–118.
- Brotans L., Thuiller W., Araújo M. B., Hirzel A. H. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability // Ecography. 2004. Vol. 27. P. 437–448.
- Burnham K. P., Anderson D. R. Kullback-Leibler information as a basis for strong inference in ecological studies // Wildlife Res. 2001. Vol. 28. P. 111–119.
- Cornwell W. K., Ackerly D. D. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California // Ecol. Monogr. 2009. Vol. 79. P. 109–126.
- Dangjin city, The statistical yearbook in 2014. Dangjin city, Chungnam Province (in Korean), 2014.
- Darveau M., Beauchesne P., Bélanger L., Huot, J., Larue P. Riparian forest strips as habitat for breeding birds in boreal forest // J. Wildlife Manage. 1995. Vol. 59. P. 67–78.
- De Alba S., Lindstrom M., Schmacher T. E., Malo D. D. Soil landscape evolution due to soil redistribution by tillage: a new conceptual model of soil catena evolution in agricultural landscapes // Catena. 2004. Vol. 58. P. 77–100.
- Dubayah R., Rich P. M. Topographic solar radiation models for GIS // Int. Journ. Geogr. Inf. Syst. 1995. Vol. 9. P. 405–419.
- Eldridge M. D. B., Pearson D. J. Black-footed rock wallaby *Petrogale lateralis* // The Mammals of Australia. 3rd edition / eds. S. M. Van Dyck, R. Strahan. Sydney: New Holland, 2008. P. 376–380.
- Elith J., Burgman M. A. Habitat models for population viability analysis // Population viability in plants / eds. C. A. Brigham, M. W. Schwanz. Springer, 2003. P. 203–235.
- Elith J., Graham C. H., Anderson R. P., Dudik M., Ferrier S., Guisan A., Hijmans R. J., Huettmann F., Leathwick J. R., Lehmann A., Li, J., Lohmann L. G., Loiselle B. A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J. M., Peterson A. T., Phillips S. J., Richardson K., Scachetti-Pereira R., Schapire R. E., Soberón J., Williams S., Wisz M. S., Zimmermann N. E. Novel methods improve prediction of species' distributions from occurrence data // Ecography. 2006. Vol. 29. P. 129–151.
- Fahrig L. Effect of habitat fragmentation on the extinction threshold: A synthesis // Ecol. Appl. 2002. Vol. 12. P. 346–353.

- Ferrier S., Watson G., Pearce J., Drielsma M. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modeling // *Biodivers. Conserv.* 2002. Vol. 11. P. 2275–2307.
- Ferrier S., Guisan A. Spatial modelling of biodiversity at the community level // *J. Appl. Ecol.* 2006. Vol. 43. P. 393–404.
- Franklin J., McCullough P., Gray C. Terrain variables used for predictive mapping of vegetation communities in southern California // *Terrain Analysis: Principles and Applications* / eds. J. P. Wilson, J. C. Galant. Wiley, 2000.
- Franklin J. Mapping species distributions: spatial inference and prediction. N.Y.: Cambridge University Press, 2010. P. 123–124.
- Friston K. J., Holmes A. P., Worsley K. J., Poline J. B., Frith C. D., Frackowiak R. S. J. Statistical Parametric Maps in functional imaging: A general linear approach // *Hum. Brain. Mapp.* 1995. Vol. 2. P. 189–210.
- Gibbons J. W. Terrestrial habitat: A vital component for herpetofauna of isolated wetlands // *Wetlands.* 2003. Vol. 23. P. 630–635.
- Gormley A. M., Forsyth D. M., Griffioen P., Lindeman M., Ramsey D. S., Scroggie M. P., Woodford L. Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species // *J. Appl. Ecol.* 2011. Vol. 48. P. 25–34.
- Grillet P., Cheylan M., Thirion J., Doré F., Bonnet X., Dauge C., Chollet S., Marchand M. A. Rabbit burrows or artificial refuges are a critical habitat component for the threatened lizard, *Timon lepidus* (Asuria, Lacertidae) // *Biodivers. Conserv.* 2010. Vol. 19. P. 2039–2051.
- Grimbacher P. S., Catterall C. P. How much do site age, habitat structure and spatial isolation influence the restoration of rainforest beetle species assemblages? // *Ibid.* 2007. Vol. 135. P. 107–118.
- Guisan A., Zimmermann N. E. Predictive habitat distribution models in ecology // *Ecol. Model.* 2000. Vol. 135. P. 147–186.
- Guisan A., Thuiller W. Predicting species distribution: offering more than simple habitat models // *Ecol. Lett.* 2005. Vol. 8. P. 993–1009.
- Hastie T., Tibshirani R. *Generalized Additive Models.* Chapman and Hall, 1990.
- Hernandez P. A., Graham C. H., Master L. L., Albert D. L. The effect of sample size and species characteristics on performance of different species distribution modeling methods // *Ecography.* 2006. Vol. 29. P. 773–785.
- Kim J. Y., Seo C. W., Kwon H. S., Ryu J. E., Kim, M. J. A Study on the Species Distribution Modeling using National Ecosystem Survey Data // *Environ. Impact Asses.* 2012. Vol. 21. P. 593–607.
- Koenig W. D. Spatial autocorrelation of ecological phenomena // *Trends Ecol. Evol.* 1999. Vol. 14. P. 22–26.
- Kullback S., Leibler R. A. On information and sufficiency // *Ann. Math. Statist.* 1951. Vol. 22. P. 79–86.
- Lawton J. H., Woodroffe G. L. Habitat and the distribution of water voles: Why are there gaps in a species' range? // *J. Anim. Ecol.* 1991. Vol. 60. P. 79–91.
- Lehmann A., Overton J. M., Leathwick J. R. GRASP: generalized regression analysis and spatial prediction // *Ecol. Model.* 2002. Vol. 157. P. 189–207.
- Legendre P. Spatial autocorrelation: trouble or new paradigm? // *Ecology.* 1993. Vol. 74. P. 1659–1673.
- Macdonald D. W., Tew T. E., Tod I. A. The ecology of weasels *Mustela nivalis* on mixed farmland in southern England // *Biologia.* 2004. Vol. 59. P. 235–241.
- Mace R. D., Waller J. S., Manley T. L., Ake K., Wittinger W. T. Landscape evaluation of grizzly bear habitat in western Montana // *Conserv. Biol.* 1999. Vol. 13. P. 367–377.
- MacKenzie D. I., Royle J. A. Designing occupancy studies: general advice and allocating survey effort // *J. Appl. Ecol.* 2005. Vol. 42. P. 1105–1114.
- Manel S., Dias J. M., Ormerod S. J. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird // *Ecol. Model.* 1999. Vol. 120. P. 337–347.
- McCollin D. Forest edges and habitat selection in birds: a functional approach // *Ecography.* 1998. Vol. 21. P. 247–260.
- McCullagh P., Nelder J. A. *Generalized Linear Models.* L.: Chapman & Hall, 1989.
- Papiernik S. K., Schumacher T. E., Lobb D. A., Lindstrom M. J., Liese M. L., Eynard A., Schumacher J. A. Soil properties and productivity as affected by topsoil movement within an eroded landform // *Soil & Tillage Res.* 2009. Vol. 102. P. 67–77.
- Patz J. A., Graczyk T. K., Geller N., Vittor A. Y. Effects of environmental change on emerging parasitic diseases // *Int. Journ. Parasitol.* 2000. Vol. 30. P. 1395–1405.
- Phillips S. J., Anderson R. P., Schapire R. E. Maximum entropy modeling of species geographic distributions // *Ecol. model.* 2006. Vol. 190. P. 231–259.
- Rangel T. F., Diniz-Filho J. A. F., Bini L. M. SAM: a comprehensive application for Spatial Analysis in Macroecology // *Ecography.* 2010. Vol. 33. P. 46–50.
- Rhim S. J., Lee W. S. Influence of forest fragmentation on the winter abundance of mammals in Mt. Chirisan National Park, South Korea // *J. Wildlife Manage.* 2007. Vol. 71. P. 1404–1408.
- Sakaki H., Ono Y. Habitat use and selection of the siberian weasel *Mustela sibirica* coreana during the non-mating season // *J. Mamm. Soc. Japan.* 1994. Vol. 19. P. 21–32.
- Scott J. M., Heglund P. J., Samson F., Haufler J., Morrison M., Raphael M., Wall B. *Predicting Species Occurrences: Issues of Accuracy and Scale.* Covelo: Island Press, 2002. P. 868.
- Thuiller W. BIOMOD – Optimizing Predictions of Species Distributions and Projecting Potential Future

- Shifts under Global Change // *Global Change Biol.* 2003. Vol. 9. P. 1353–1362.
- Tratalos J., Fulle R. A., Warren P. H., Davies R. G., Gaston K. J. Urban form, biodiversity potential and ecosystem services // *Landscape Urban Plan.* 2007. Vol. 83. P. 308–317.
- Vos C. C., Chardon J. P. Effects of habitat fragmentation and road density on the distribution pattern of the moor frog *Rana arvalis* // *J. Appl. Ecol.* 1998. Vol. 35. P. 44–56.
- William J. R., Robert L. B. Wolves and the Ecology of Fear: Can Predation Risk Structure Ecosystem? // *BioScience.* 2004. Vol. 54. P. 755–766.
- Wisz M. S., Hijmans R. J., Li J., Peterson A. T., Graham C. H., Guisan A. NCEAS Predicting Species Distributions Working Group, Effects of sample size on the performance of species distribution models // *Divers Distrib.* 2008. Vol. 14. P. 763–773.
- Yackulic C. B., Chandler R., Zipkin E. F., Royle J. A., Nichols J. D., Grant E. H. C., Veran S. Presence-only modelling using MAXENT: When can we trust the inferences? // *Methods Ecol. Evol.* 2012. Vol. 4. P. 236–243.
- Zhang E. D., Teng L. W., Wu Y. B. Habitat suitability evaluation for the Chinese water deer *Hydropotes inermis* in Yancheng Nature Reserve, China // *Acta Theriologica Sinica.* 2006. Vol. 26. P. 368–372.