

Article

# Habitat Potential Mapping of Marten (*Martes flavigula*) and Leopard Cat (*Prionailurus bengalensis*) in South Korea Using Artificial Neural Network Machine Learning

Saro Lee <sup>1,2</sup> , Sunmin Lee <sup>3,4</sup> , Wonkyong Song <sup>5</sup> and Moun-Jin Lee <sup>4,\*</sup> 

<sup>1</sup> Geological Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124, Gwahak-ro Yuseong-gu, Daejeon 34132, Korea; leesaro@kigam.re.kr

<sup>2</sup> Korea University of Science and Technology, 217 Gajeong-ro Yuseong-gu, Daejeon 34113, Korea

<sup>3</sup> Department of Geoinformatics, University of Seoul, 163 Seoulsiripdaero, Dongdaemun-gu, Seoul 02504, Korea; smlee@kei.re.kr

<sup>4</sup> Center for Environmental Assessment Monitoring, Environmental Assessment Group, Korea Environment Institute (KEI), 370 Sicheong-daero, Sejong 30147, Korea

<sup>5</sup> Department of Landscape Architecture, Dankook University, 119 Dandae-ro, Dongnam-gu, Cheonan-si, Chungnam 31116, Korea; wksong@dankook.ac.kr

\* Correspondence: leemj@kei.re.kr; Tel.: +82-44-415-7314

Received: 15 July 2017; Accepted: 31 August 2017; Published: 5 September 2017

**Abstract:** This study developed habitat potential maps for the marten (*Martes flavigula*) and leopard cat (*Prionailurus bengalensis*) in South Korea. Both species are registered on the Red List of the International Union for Conservation of Nature, which means that they need to be managed properly. Various factors influencing the habitat distributions of the marten and leopard were identified to create habitat potential maps, including elevation, slope, timber type and age, land cover, and distances from a forest stand, road, or drainage. A spatial database for each species was constructed by preprocessing Geographic Information System (GIS) data, and the spatial relationship between the distribution of leopard cats and environmental factors was analyzed using an artificial neural network (ANN) model. This process used half of the existing habitat location data for the marten and leopard cat for training. Habitat potential maps were then created considering the relationships. Using the remaining half of the habitat location data for each species, the model was validated. The results of the model were relatively successful, predicting approximately 85% for the marten and approximately 87% for the leopard cat. Therefore, the habitat potential maps can be used for monitoring the habitats of both species and managing these habitats effectively.

**Keywords:** habitat mapping; marten; leopard cat; ANN; South Korea

## 1. Introduction

Medium-sized predators control not only herbivorous animals but also intermediate predators, and have a combined effect on the overall food chain. Apex predators generally have low population densities and wide ranges of activities, and are vulnerable to local extinction due to habitat damage and disconnection [1]. In South Korea, members of the order Carnivora, such as tigers, leopards, and wolves, have become extinct or lost their ecological functions in natural ecosystems over the past century. As a result, the importance of the remaining carnivorous animals, such as martens, leopard cats, and otters, is increasing.

Martens (*Martes* sp.) are mammals with wide distribution worldwide, from subtropical to sub-Antarctic habitats. They are generally characterized by a larger sphere of activity than other

mammals [2], and have roles as forest ecosystem indicator species sensitive to habitat disturbances and habitat fragmentation [3]. Therefore, the marten in Korea, the yellow-throated marten (*Martes flavigula*), is considered a second-grade endangered species of wild fauna and flora by the Ministry of Environment, South Korea, and belongs to Annex II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Studies on marten habitats have been conducted mainly in tropical and subtropical areas such as Thailand [4], and there is insufficient information on marten habitats on the Korean Peninsula.

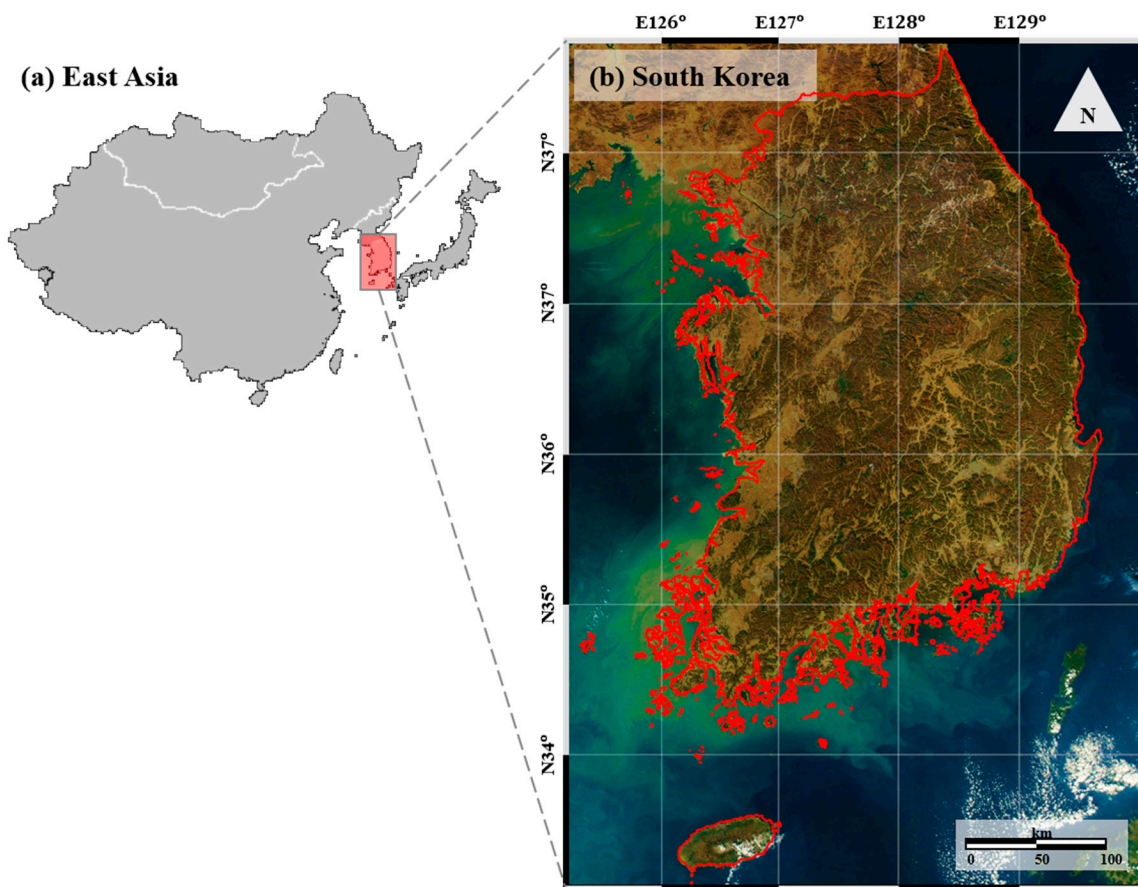
The leopard cat (*Prionailurus bengalensis*) is the only wild feline carnivore in South Korea, which also belongs to the second-grade endangered species of wild fauna and flora of the Ministry of Environment and the CITES Annex II list. Leopard cats are distributed within a wide range, from deep forests to rural and coastal areas in Korea. Moreover, they maintain relatively stable populations, due to their adaptability to inhabiting various habitats. It has been reported that leopard cats are distributed widely in forests and rural areas [5], but there have been few quantitative studies on the environmental spatial characteristics of their habitats. In particular, there is a lack of research on their distribution throughout South Korea and analysis of their habitat characteristics.

Machine learning is designed to process new data and predict results by learning patterns through training based on consistent patterns among variables in a dataset [6,7]. Recently, the amount of plant- and animal-related data has rapidly increased, and many studies using machine learning models have been conducted [8,9]. Depending on the approach applied, machine learning can use various algorithms, such as decision tree, neural network, and support vector machine, and it can be divided into supervised learning and unsupervised learning depending on the presence or absence of training data. Thus, many studies have been conducted to predict the distribution of plants and animals using machine learning techniques [10,11]. In addition, Geographic Information System (GIS) platforms have been used as a useful tool to model the spatial relationships between specific events and related factors [12–14].

Thus, recent studies have used GIS to indicate the distribution of habitats of various species. Studies on mapping and quantifying mammalian habitats have been conducted through GIS-based models [15,16]. Among mammals, studies on martens and leopard cats have been conducted mainly in Asia. Studies of the genetic structure and mitochondrial genome of the marten and the leopard cat have also been performed [17,18]. In particular, the complete mitochondrial genome has been analyzed for martens in Korea [19]. Meanwhile, studies on the movements and activity patterns of leopard cats have been conducted in China [20,21]. One study predicted the distribution of leopard cats in Borneo [22], and another similar study predicting the habitat potential of leopard cats was conducted in South Korea using GIS [23]. Habitat use and activity patterns of martens were also analyzed in central and northern Thailand [4]. Typically, the habitat distributions of badgers [24,25], leopard cats [26], and bears [27] have been mapped using logistic regression models. In recent years, research has been conducted on the creation of habitat distribution maps of leopard cats in South Korea through probabilistic and statistical models [23]. Supporting this, other probabilistic models have been used to map European bison habitats [28], and statistical models have been applied to the habitat distribution of bats [29,30]. Moreover, habitat mapping of water birds for conservation plans was conducted in the Hamun wetland [31]. In Canada, a high-resolution habitat map of Atlantic wolffish was also created [32]. However, artificial neural network (ANN) modeling has not been applied to analyzing the habitat of martens or leopard cats in South Korea. Therefore, the purposes of this study were to map the distribution of martens and leopard cats in South Korea using ANN modeling and to clarify the relationship between the habitat distributions and various environmental factors.

South Korea is located in East Asia between latitudes 33° and 39° N and longitudes 124° and 130° E, including all of its islands, and occupies the southern part of the Korean Peninsula (Figure 1); the total area of the country is approximately 100,032 km<sup>2</sup> [33]. The boundary of the study area in this study is marked with a red line in the Figure 1. South Korea can be divided into four regions: a western region of broad coastal plains, an eastern region of high mountain ranges, a southwestern mountainous

region, and a southeastern region of a broad basin. Mountains cover about 70% of South Korea, and the country is surrounded by the Yellow Sea to the west, the East Sea to the east, and the Korea Strait and the East China Sea to the south. The north of the country is bordered by the Democratic People's Republic of Korea, via the demilitarized zone; therefore, it is difficult to support the return of extirpated terrestrial animals in South Korea, since their path is blocked in all directions. To support conservation efforts, habitat potential maps should be created to manage habitats before extirpation, including the selection of development areas. Therefore, in this study, ANN modeling was applied to generate habitat potential maps for martens and leopard cats. The weights of factors related to the habitat potentials were calculated with the model, and the results were validated to ensure the reliability of the maps.



**Figure 1.** Study area: (a) South Korea in East Asia; and (b) South Korea.

## 2. Data

### 2.1. Habitat Survey

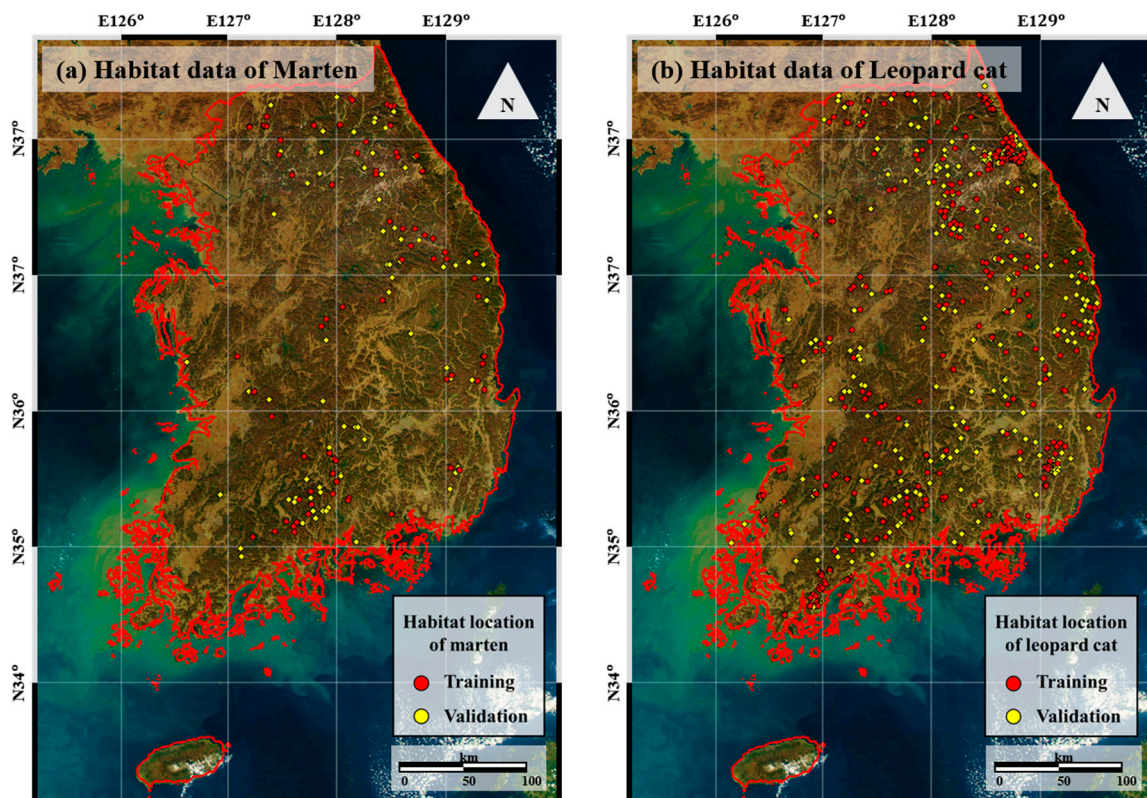
This study used data from the Second National Survey on Natural Environment in South Korea, which included the identification of habitats of endangered species of wild animals and plants. From 1997 to 2003, the National Institute for Environmental Studies conducted a survey of species appearances and spatial distributions of wild animals by experts from various research institutes. Surveys were performed from February to October every year for mammals to consider changes in seasonal patterns. The survey methods were based on field observations, including direct observations, community surveys, tracking, feces, and footprints, to investigate the species occurrence and spatial distribution of wild animals. The locations observed or detected by the traces were geocoded via Global Positioning System and were composed of GIS data.



The marten has the second largest sphere of action among mammals in the Korean Peninsula after the Asiatic black bear. Martens prefer broadleaf forests and mainly inhabit mature forests, such as fourth-grade forests that have  $\geq 50\%$  occupancy rates of 31–40-year-old trees with diameters of  $\geq 30$  cm. Martens inhabiting the Jirisan area occur at a density of 1–1.6 per 10 km<sup>2</sup>, and are diurnal animals with a wide range of behaviors, moving an average of  $11.2 \pm 5.4$  km per day. Martens mainly hunt Eurasian red squirrels, rats, and other rodents as food, but they also hunt hares, young roe deer, and wild boars. They also consume tree fruit, such as those from the lotus persimmon tree. In winter, they often hunt for food in shrubs on the edge of forests and frequently cross two-lane roads in their radius of action [34].

Leopard cats have activity areas of  $3.69 \pm 1.34$  km<sup>2</sup>, and the core space of the species is estimated to be  $0.64 \pm 0.47$  km<sup>2</sup> [5]. Leopard cats prefer forests, as well as adjacent grassland and agricultural land, as habitats. They have a high preference for inland wetlands, such as wild grassland along riverbanks. Owing to their behavioral radius and nocturnality of movement in a variety of regions, they suffer a high frequency of road kills on roads adjacent to forests and rivers. Rodents, birds, and small mammals are their main sources of food, and forest ridges and valleys are used as main transport routes by leopard cats.

In the Second National Survey on Natural Environment, martens were found at 156 points, and leopard cats were found at 630 points in the study area (Figure 2). In this study, the observed data were divided randomly in half for model training and validation. This confirmed that the distribution of martens is mainly limited to large forest areas, including the Mt. Baekdu range. Meanwhile, leopard cats inhabit a wide range of emergence sites, from major forests to rivers and agricultural areas.



**Figure 2.** Data obtained from field observations of (a) Martens; and (b) Leopard cats.

## 2.2. Habitat-Related Factors

Habitat distributions of wild animals are influenced by the combined impacts of various factors. Environmental factors such as topographic characteristics, forest properties [22], and distance from

essential factors to support life, such as water, influence the habitat distributions of mammals, including martens and leopard cats [22]. In this study, ground elevation, slope gradient, slope aspect, timber type, timber age, land cover, and distance from road, water, and forest were selected as factors influencing the habitat distributions of marten and leopard cat (Table 1 and Figure 3). The attribute information of the timber type of forest map is shown separately in Table 2, due to the limitation of the figure size. The determination and collection of factors were a fundamental component of mapping the potential habitats of these species. The factors were collected from nationally generated thematic maps and field investigations, such as those described in Section 2.1.

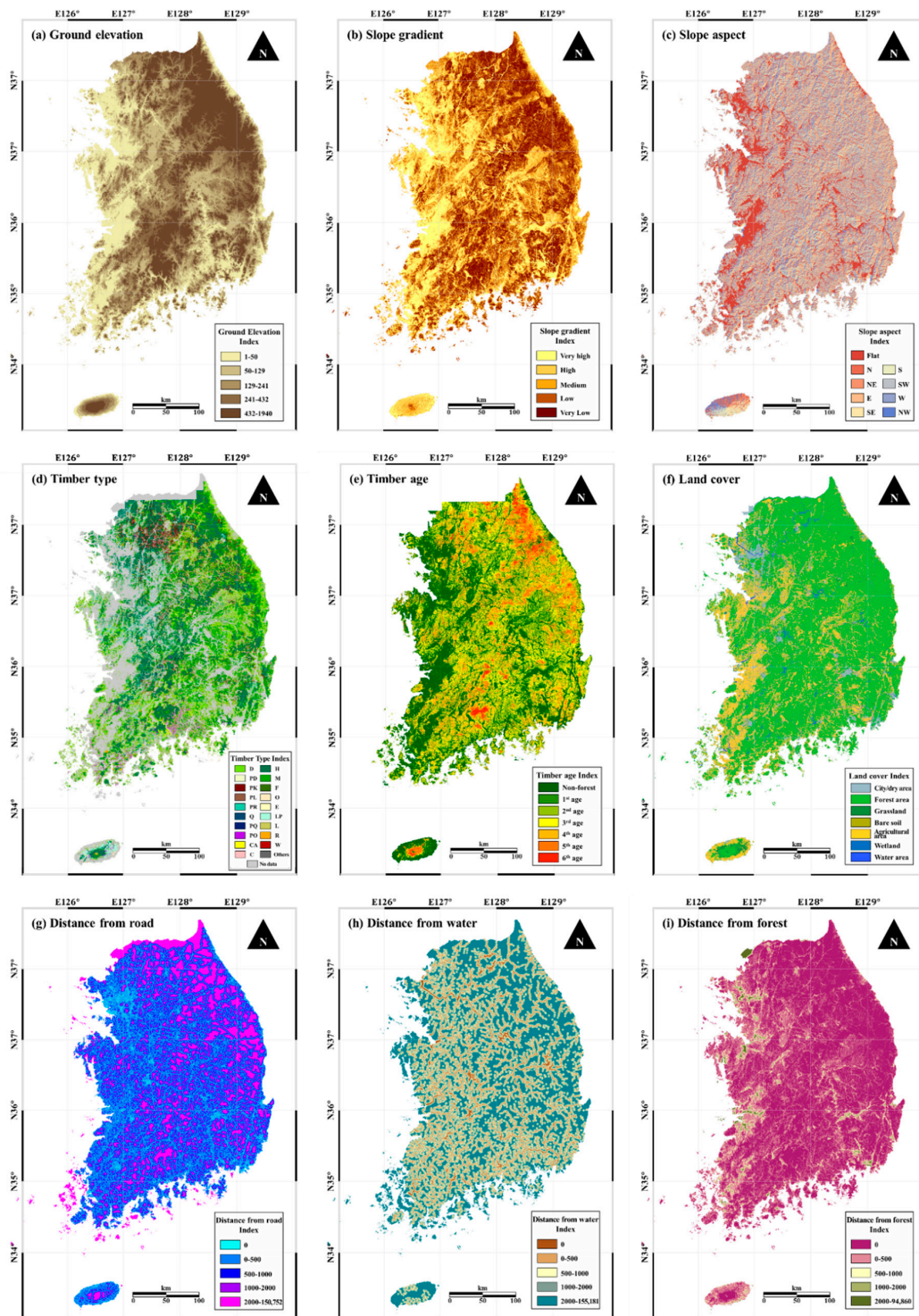
**Table 1.** Data layers related to marten and leopard cat habitat.

Original Data	Factors	Data Type	Scale
Habitat	Marten Leopard cat	Point	-
Topographical map <sup>a</sup>	Ground elevation (m) Slope gradient (°) Slope aspect	GRID	1:5000
Forest map <sup>b</sup>	Timber type Timber age	Polygon	1:25,000
Land cover map <sup>c</sup>	Land cover Distance from road (m) Distance from water (m) Distance from forest (m)	Polygon	1:25,000

<sup>a</sup> The digital topographic map by National Geographic Information Institute (NGII); <sup>b</sup> The forest map published by Korea Forest Service (KFS); <sup>c</sup> The land use map offered by the Korea Ministry of Environment.

A digital elevation model (DEM) was produced by generating a triangulated irregular network and digitizing the contours in 100-m intervals from topographical maps published by the National Geographic Information Institute. The slope gradient and slope aspect calculations were performed in ArcGIS 10.3. Thematic maps of forest and land cover were prepared at a 1:25,000 scale in a polygon format. The forest map was provided by the Korea Forest Service, and the timber type and age were prepared from this map. The land cover map was generated and published by the Ministry of Environment, South Korea. Land cover, distance from road, water, and forest were also calculated using ArcGIS.

All factors were converted into a 100 m × 100 m grid format to apply the ANN model. The row and column sizes of the study area were 6056 and 3533 cells, and the study area consisted of a total of 9,952,165 cells, except in the no data region, based on the DEM. The marten and leopard cat habitat data were converted from point formats into grid formats. Half of both the marten and the leopard cat habitat data were selected via random sampling as input data. The other half of the data were used for validation. Ultimately, all factors were collated into one input dataset with randomly extracted habitat sample data for both marten and leopard cat.



**Figure 3.** Factors related to the habitat potentials of marten and leopard cat: (a) ground elevation, (b) slope gradient, (c) slope aspect, (d) timber type, (e) timber age, (f) land cover, (g) distance from road, (h) distance from water, and (i) distance from forest.



**Table 2.** Attribute information of timber type from forest map.

	Code	Forest Type		Code	Forest Type
Forest species	D	Pinus densiflora Forests	Forest physiognomy	C	Conifer mixed forest
	PD	Pinus densiflora artificial forest		H	Broadleaved forest
	PK	Pinus koraiensis forest		M	Mixed forest of soft and hardwood
	PL	Larch	Dentuded area	F	Cut-over area
	PR	Pinus rigida forest		O	Non-stocked forest land
	Q	Oak forest		E	Dentuded land
	PQ	Oak artificial forest		LP	Grassland
	PO	Poplar forest	Left-over area	L	Farmland
	CA	Chestnut artificial forest		R	Left-over area
				W	Water

### 3. Methods

The machine learning technique determines the algorithm patterns using the sampled training data and derives the results through learned patterns. In other words, it is possible to learn from results that are already known according to information from the samples. Habitats are influenced by various environmental factors based on the nature of animals; therefore, even if there are no animals at the present time, the possibility of migration into an area with a similar environment exists at any time. Therefore, the habitat pattern can be determined by training based on habitat data using the machine learning method, and the habitat potential of an entire study area can be created. Among various machine learning techniques, habitat potential mapping of martens and leopard cats was conducted using an ANN in this study.

Interest in ANNs has increased recently. ANN is a highly sophisticated modeling method that can model complex functions such as environmental problems or social issues with large numbers of variables. Mimicking the human biological neural system, an ANN adjusts the weight from the basic unit of perceptron between the input and output data, minimizing the error of the result. In addition to these strengths, ANNs have been used for prediction or classification in a remarkable range of fields, such as engineering, physics, geology, and environmental science. In particular, neural networks can be applied to nonlinear problems of dimensionality and can be used to identify spatial relationships between observation location data and influencing factors.

Multi-layer perceptron is one of the most popular ANN models, and was created by [35]. This network, which allows for the regression of nonlinear data, was used in this study with a hidden layer between the input layers of habitat-related factors and the output layer. In this study, a back-propagation algorithm was used for neural network training. In back propagation, the algorithm calculates the gradient vector of the error surface and distance from the current point. The error can be decreased when the point moves a short distance from the output layer to other layers. The process iterates through a number of epochs, submitting the training data and calculating the error by comparing the target and output. The weights can be corrected through the error from the iteration and surface gradient. The initial network is randomly set, and the stopping point is determined by a set number of epochs or a user-selected stopping point. The following equation indicates the weight-updating architecture of the back-propagation algorithm.

$$\omega_k^{up} = \omega_k - \eta * \frac{\partial E_{total}}{\omega_k} \tag{1}$$

where  $\omega_{k_0}$  is the initial value of the weight  $\omega_k$ , which is given randomly, and  $\omega_k$  is the updated weight for connection  $k$ .  $\eta$  is the learning rate that determines the step size and is typically chosen experimentally.  $E_{total}$  is the error of the output.

$$E_{total} = \frac{1}{2} \sum_j^n (T_j - out_{o_j})^2 \tag{2}$$

$T$  is the expected output value according to the input data [36] and  $j$  is the number assigned to the node in each layer.  $n$  is the number of output layers.  $E_{total}$  is minimized through the updated weight from Equation (16). To detect the minimum point, differentiation of  $E_{total}$  is necessary. Thus, an activation function is needed for conversion from the result of real number  $x$  to a range of values between 0 and 1 when the value is passed through the next layer or when the final result is expected. The unipolar sigmoid function was used for the activation function in this study. The equation of the unipolar sigmoid function is as follows.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

In this study, the ANN was supported by MATLAB software. Using the training data of randomly extracted habitat sample data, the slope data of the flat area were used as non-habitat area for training, and the input data of the study area were rearranged between 0.1 and 0.9. The weights of the input layers were calculated using the back-propagation algorithm of the ANN with a  $9 \times 18 \times 1$  structure of the networks. The initial weight was set randomly as described previously, and the learning rate and number of epochs were set to 0.01 and 1000, respectively. The root-mean-square error, the criterion of the stopping point of decreasing error, was set to 0.01.

The process of potential mapping for martens and leopard cats can be explained in three steps, as shown in Figure 4. First, a spatial database was constructed with nine potential habitat-influencing factors, including the observation points of martens and leopard cats. Half of the marten and leopard distribution data selected by random sampling were used as training sets. The locations of observations of martens and leopard cats were confirmed in field surveys for the application and validation of habitat potential models described in Section 2.2 (Figure 2). The observation location data of martens and leopard cats were set as dependent variables. Second, the ANN model was applied to a spatial database to map the habitat potentials. Nine factors that were considered to affect the marten and leopard cat habitats were set as independent variables. Elevation, slope, and aspect derived from topographic maps, timber type and timber age from forest maps, and land cover, distance from road, distance from water, and distance from forest from land cover map were included in the spatial database with half of the randomly sampled marten and leopard cat habitat data. The ANN model was applied after constructing the spatial database. To apply the model, MATLAB programming was used with the GIS data. The resulting maps of the habitat potentials were validated using a receiver operating characteristic (ROC) curve. Finally, validation of the marten and leopard cat habitat potential maps was performed using the remaining marten and leopard cat distribution data that were not included in the spatial database.

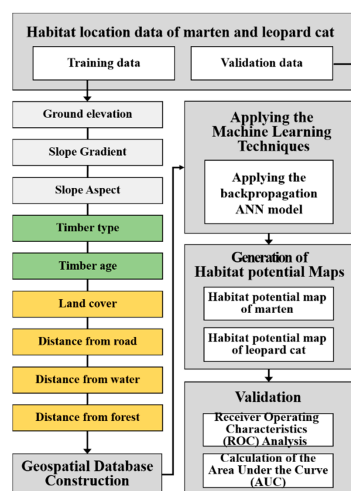


Figure 4. Flow chart of the steps used in this study.



## 4. Results

### 4.1. Weight of Related Factors and Habitat Potential Mapping

The weights of the related factors between the layers acquired during the ANN model training process are shown in Table 3. The weight values show the contribution of each weight to the ANN model. Since the weights were initialized during the first training, the results could differ. This study attempted to obtain similar results by performing the calculation several times. The weight calculation was repeated for 10 epochs, with 100 cycles for each epoch, to identify the influence of the randomly extracted sample for a city/dry area (public area), non-habitat area, since most wild animals do not live in populated urban areas. No leopard cats were observed in public areas in the results of the frequency ratio analysis in a previous study [23]. In the iterations of each epoch, the weights were updated via back propagation. The standard deviation was 0.001–0.018 for martens and 0.002–0.021 for leopard cats. The ranges of standard deviations indicated that there were minimal effects of random sampling of non-habitat area on the results.

**Table 3.** Neural network weight between martens and leopard cats, and habitat-related factors.

	Marten			Leopard Cat		
	Average	Standard Deviation	Normalized Weight with Respect to Land Cover	Average	Standard Deviation	Normalized Weight with Respect to Land Cover
DEM	0.1657	0.0175	2.3877	0.1728	0.0208	2.3883
Slope gradient	0.1335	0.0102	1.9242	0.1939	0.0103	2.6790
Slope aspect	0.0901	0.0019	1.2986	0.0768	0.0019	1.0619
Timber type	0.1453	0.0107	2.0945	0.1159	0.0185	1.6014
Timber age	0.0704	0.0104	1.0146	0.1053	0.0059	1.4545
Land cover	0.0694	0.0015	1.0000	0.0724	0.0030	1.0000
Distance from road	0.1092	0.0023	1.5741	0.0817	0.0054	1.1285
Distance from water	0.1210	0.0032	1.7439	0.0854	0.0037	1.1803
Distance from forest	0.0953	0.0008	1.3735	0.0959	0.0044	1.3259

Regarding the average weights of the related factors for the habitat potential mapping of martens, DEM had the highest value (0.166), followed by timber type (0.145), while land cover showed the lowest value (0.069). The factor weights for leopard cats were similar. Slope gradient had the highest value (0.194), followed by DEM (0.173), and land cover had the lowest value (0.072).

However, for a comparison of the same standards between the two analyses, all average weight values were normalized by dividing by the smallest average weight value, land cover, for each factor for martens and leopard cats. The lowest values of 0.069 and 0.072 were normalized to 1.000. DEM showed the highest normalized weight value of 2.388 for martens and timber type was the second highest normalized factor at 2.095. Likewise, the weight analysis showed that the slope gradient and DEM presented the highest and second highest normalized weight values of 2.679 and 2.388 for leopard cats, respectively. The least important factors, land cover for both martens and leopard cats, were used as references for the normalized weights. Timber age (1.015) and slope aspect (1.299) for martens and slope aspect (1.062) and distance from road (1.129) for the leopard cats followed.

The weights calculated via iteration, shown above, were applied throughout the study area, and the final weights determined for martens and leopard cats were applied to the corresponding factor in each dataset. Figure 5 shows the habitat potential maps for martens and leopard cats. To simplify the interpretation of the habitat potential maps, the results were classified into five classes of potentials: very high (10%), high (10%), medium (20%), low (20%), and very low (40%). Based on the habitat potential maps of both martens and leopard cats, the eastern region of South Korea appeared to have high potential.

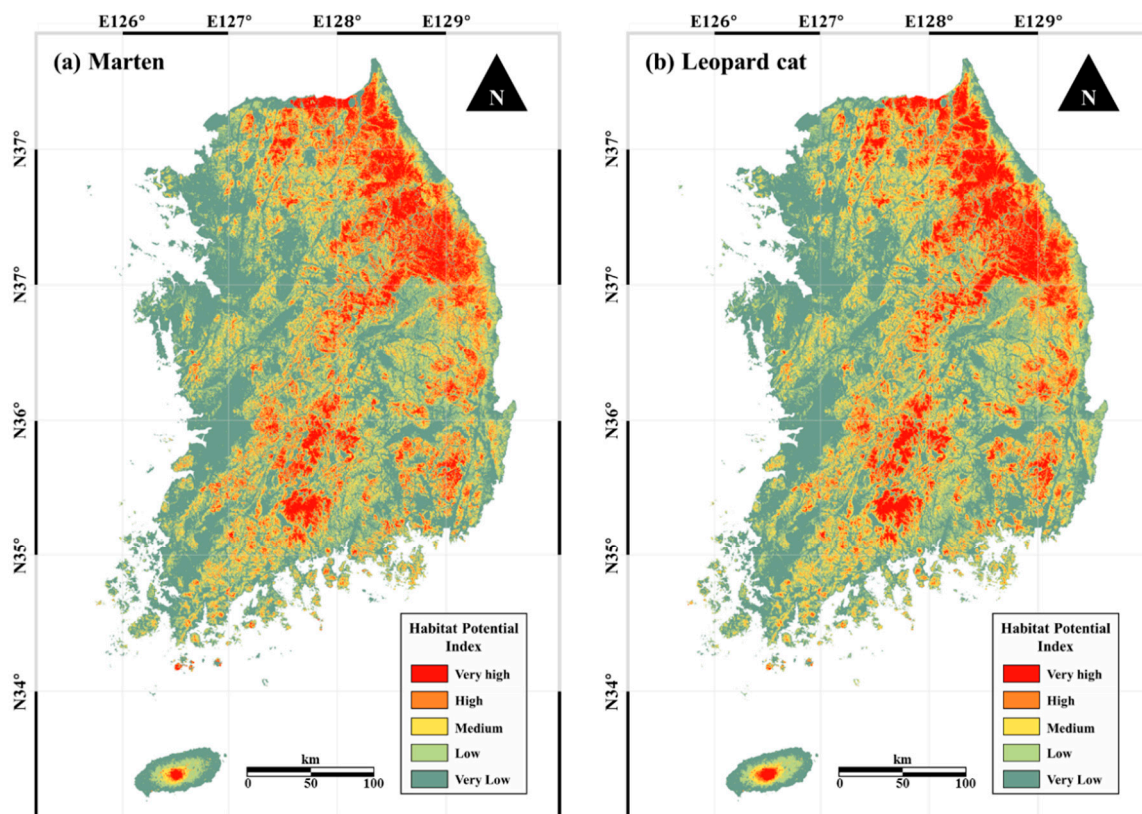


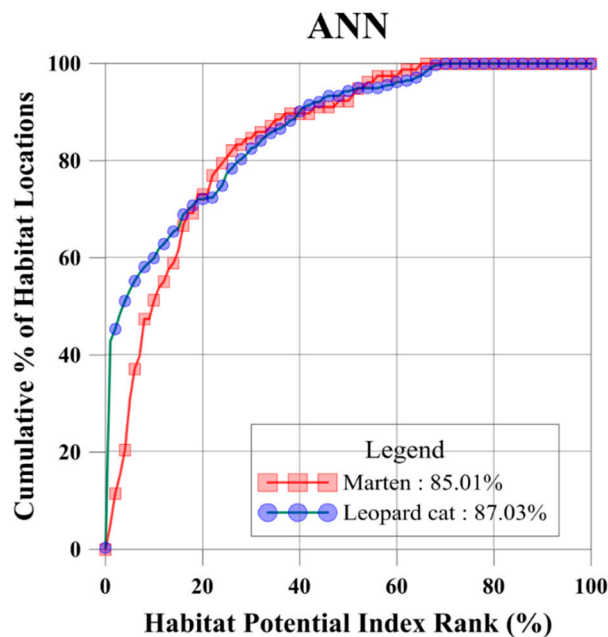
Figure 5. Habitat potential maps generated using ANN model of (a) martens; and (b) leopard cats.

#### 4.2. Validation

Habitat data can be used to validate whether the habitat potential maps were effectively predicted. Therefore, the halves of marten and leopard cat habitat data not used for training were used as the validation data. ROC curves were used to validate the habitat potential maps of martens and leopard cats generated using the ANN model. The ROC curves were generated for the marten and leopard cat habitat potential maps to compare the habitat locations of martens and leopard cats, respectively.

The ROC curve is a graphical plot showing the diagnostic capabilities of classification models with various thresholds. It is a widely used method for validation [37–40]. ROC curves can show sensitivity and specificity on the  $x$ -axis and  $y$ -axis, respectively. In this study, specificity represented the percentage of area that martens and leopard cats could inhabit, while sensitivity represented the predicted potential marten and leopard cat habitat locations. For the ROC graph, the predicted habitat potential values after applying the ANN model were sorted in descending order. The ranks of the habitat potential values were identified and the values were equally classified into 100 classes of study area.

Figure 6 shows the ROC curves of the marten and leopard cat habitat potential maps; martens showed 85.01% (0.8501) accuracy, versus 87.03% (0.8703) accuracy for leopard cats. These results are depicted as a graph of the validation rate (Figure 6), where 65% of the study area included 100% of all habitat locations for both species.



**Figure 6.** Validation results of the habitat potential maps for martens and leopard cats using the ANN model.

## 5. Discussion

Machine learning focuses on processing new data and predicting outcomes using patterns learned through training [12,13]. In other words, the most noticeable advantage of ANN modeling is that the algorithm itself can learn to be more accurate, although it requires sufficient data to support this. In South Korea, a government-affiliated organization constructed detailed country-scale thematic maps. These previously constructed data were sufficient for generalizing and correctly processing new incoming untrained data. Therefore, a spatial database was constructed based on nationally constructed data, including topographical, forest, and land use maps. The spatial database built in this process could be used in conjunction with other methods. MATLAB was used to process large amounts of data quickly, and the spatial distribution of martens and leopard cats could be confirmed through a GIS analysis. The results for leopard cats using the ANN showed approximately 5% higher accuracy than those obtained using the frequency ratio (82.15%) and logistic model (81.48%) in a previous study [23], even though the input data differed. This confirmed that the ANN model had been applied appropriately to habitat potential mapping.

The results of machine learning models, such as the ANN model applied in this study, are strongly influenced by the input data. Therefore, in this study, an input spatial database was constructed with marten and leopard cat habitat point data based on the Second National Survey on Natural Environment, and data extracted and calculated from national theme maps, which reflected regional characteristics. The impacts of factors besides the input data, such as climate or sudden natural disasters, on martens and leopard cats were assumed to be negligible. In the case of the habitat data, the generation of errors may depend on the survey method. However, the behavior radii of martens and leopard cats were within 10 km, and the size of the entire study area was considerably larger by comparison; therefore, all data were considered to be reliable.

In both species, DEM, slope gradient and timber type were the three most important factors related to habitat. The relative weight of slope gradient was higher for leopard cats than for martens, indicating a greater influence of slope gradient in the result map for leopard cats (Figure 5). The smaller radius of action of leopard cats compared to martens may have influenced this result. Martens mainly inhabit forests; therefore, the habitat potential of martens was high in relatively dense forests. In contrast,

the habitat potential of leopard cats tended to be evenly distributed in relatively low areas compared to martens, which was particularly evident on Jeju Island.

Considering the results of the habitat potential map of this study, elevation, slope gradient and timber type information of mountain areas should be considered as indices for prioritizing the protection and management of marten and leopard cat habitats. Moreover, the index should be managed with weighted indices for each factor. In the future, national natural environment survey data that are constructed and accumulated in a timely manner will enable the continuous construction and management of habitat maps for not only marten and leopard cat, but also other endangered species.

## 6. Conclusions

In this study, ANN modeling was used to predict the habitat potentials for martens and leopard cats. The factors affecting the habitat distribution of each species were selected and a spatial database was constructed with the collected data. Using the collected data and half of the field-surveyed habitat data, the habitat potentials of martens and leopard cats were predicted using the ANN model. The predicted maps were validated with the other half of the habitat data not used for training.

The first- to third-highest normalized weights with respect to the lowest average weight for both martens and leopard cats (land cover) were DEM, slope gradient and timber type. This indicates that factors related to DEM are positively correlated with the habitat location of martens and leopard cats. In contrast, land cover showed the lowest normalized weights for both martens and leopard cats. Subsequently, martens had lower weights for timber age and slope aspect, whereas leopard cats had lower weights for slope aspect and distance from road. Since both species are mammals with large activity radii, and steep slopes are considered an important factor, the weight value for distance from mountains was relatively low.

In South Korea, martens and leopard cats exhibited similar habitat patterns, with high habitat potentials in eastern mountainous areas, except along beach lines and southern areas. Likewise, coastal areas beyond the eastern range and western plains of the study area showed very low habitat potentials for both species. Nearly all areas with low habitat potentials were lowlands, coastal areas, and non-forest areas. These trends are likely similar even though the habitat characteristics of martens and leopard cats differ because of the large size of the study area.

The results of the marten and leopard cat habitat potential maps were validated with the half of the marten and leopard cat habitat data not used for modeling. The results of the validation showed accuracy of 85.01% for martens and 87.03% for leopard cats, both of which were satisfactory (>85%). These results could lead to more informed decisions for wildlife management planning and land development planning in areas with high habitat potentials.

Owing to the characteristics of the machine learning model, input data accuracy is very important. However, it is difficult to determine the exact habitat location of mammals with large activity radii, such as martens and leopard cats. Since inaccurate location data could result in difficulties in performing spatial analyses, reliable habitat data should be used when research is conducted for smaller administrative units. Quantitative evaluation of the ecological consequences of a wide range of spatial data could be performed using GIS. Integrating several characteristics representing the habitat potential of mammals is an important aspect of ecological management research.

This study identified habitat-related factors for martens and leopard cats. The methodology applied in this study could also be used to generate a time-series of habitat maps, and marten and leopard cat habitat data could be used, since the habitat-related data will be additionally and continuously constructed during the next national natural environment survey. In addition to martens and leopard cats, this method can also be applied to habitat mapping of other mammalian species examined in national natural environment surveys. Moreover, the habitat potential mapping results can be used as a basis for determining the locations of monitoring sites or creating protection plans for mammalian species. However, more studies are essential to generalize the habitat-related factors that affect the habitat characteristics for individual mammalian species.



**Acknowledgments:** This research was conducted by Korea Environment Institute (KEI) with support of a grant (16CTAP-C114629-01) from Technology Advancement Research Program (TARP) funded by Ministry of Land, Infrastructure and Transport of Korean government. This research was supported by the Basic Research Project of the Korea Institute of Geoscience and Mineral Resources (KIGAM) funded by the Ministry of Science and ICT. This research (NRF-2016K1A3A1A09915721) was supported by Science and Technology Internationalization Project through National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (MEST).

**Author Contributions:** Saro Lee organized the paperwork and constructed the input database. Sunmin Lee designed and performed the experiments. Wonkyong Song interpreted the input data and this resulted in the habitat aspect. Moungh-Jin Lee suggested the idea and collected the data. All authors contributed to the writing of each part.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ripple, W.J.; Estes, J.A.; Beschta, R.L.; Wilmers, C.C.; Ritchie, E.G.; Hebblewhite, M.; Berger, J.; Elmhagen, B.; Letnic, M.; Nelson, M.P. Status and ecological effects of the world's largest carnivores. *Science* **2014**, *343*, 1241484. [[CrossRef](#)] [[PubMed](#)]
- Taylor, S.L.; Buskirk, S.W. Forest microenvironments and resting energetics of the American marten *Martes americana*. *Ecography* **1994**, *17*, 249–256. [[CrossRef](#)]
- Harrison, D.J.; Fuller, A.K.; Proulx, G. *Martens and Fishers (Martes) in Human-Altered Environments: An International Perspective*; Springer Science & Business Media: Berlin, Germany, 2004.
- Grassman, L.I., Jr.; Tewes, M.E.; Silvy, N.J. Ranging, habitat use and activity patterns of binturong arcticis binturong and yellow-throated marten *Martes flavigula* in north-central Thailand. *Wildl. Biol.* **2005**, *11*, 49–57. [[CrossRef](#)]
- Park, C.-H.; Choi, T.-Y.; Kwon, H.-S.; Woo, D.-G. Habitat selection and management of the leopard cat (*Prionailurus bengalensis*) in a rural area of Korea. *Korean J. Environ. Ecol.* **2012**, *26*, 322–332.
- Ramcharan, A.; Hengl, T.; Nauman, T.; Brungard, C.; Waltman, S.; Wills, S.; Thompson, J. Soil property and class maps of the conterminous us at 100 meter spatial resolution based on a compilation of national soil point observations and machine learning. *arXiv*, **2017**, arXiv:1705.08323.
- Veronesi, F.; Korfiati, A.; Buffat, R.; Raubal, M. Assessing accuracy and geographical transferability of machine learning algorithms for wind speed modelling. In Proceedings of the International Conference on Geographic Information Science, Wageningen, The Netherlands, 9–12 May 2017; Springer: Berlin, Germany, 2017; pp. 297–310.
- Folmer, E.O.; van Beusekom, J.E.; Dolch, T.; Gräwe, U.; Katwijk, M.M.; Kolbe, K.; Philippart, C.J. Consensus forecasting of intertidal seagrass habitat in the wadden sea. *J. Appl. Ecol.* **2016**, *53*, 1800–1813. [[CrossRef](#)]
- Sahragard, H.P.; Chahouki, M.A.Z. Comparison of logistic regression and machine learning techniques in prediction of habitat distribution of plant species. *Range Manag. Agrofor.* **2016**, *37*, 21–26.
- Fox, C.; Huettmann, F.; Harvey, G.; Morgan, K.; Robinson, J.; Williams, R.; Paquet, P. Predictions from machine learning ensembles: Marine bird distribution and density on Canada's pacific coast. *Mar. Ecol. Prog. Ser.* **2017**, *566*, 199–216. [[CrossRef](#)]
- Wieland, R.; Kerkow, A.; Früh, L.; Kampen, H.; Walther, D. Automated feature selection for a machine learning approach toward modeling a mosquito distribution. *Ecol. Model.* **2017**, *352*, 108–112. [[CrossRef](#)]
- Anderson, D.J.; Rojas, L.F.; Watson, S.; Knelson, L.P.; Pruitt, S.; Lewis, S.S.; Moehring, R.W.; Bennett, E.E.S.; Weber, D.J.; Chen, L.F. Identification of novel risk factors for community-acquired clostridium difficile infection using spatial statistics and geographic information system analyses. *PLoS ONE* **2017**, *12*, e0176285. [[CrossRef](#)] [[PubMed](#)]
- Bui, D.T.; Bui, Q.-T.; Nguyen, Q.-P.; Pradhan, B.; Nampak, H.; Trinh, P.T. A hybrid artificial intelligence approach using gis-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agric. For. Meteorol.* **2017**, *233*, 32–44.
- Rahimi, S.; Martin, M.J.; Obeysekere, E.; Hellmann, D.; Liu, X.; Andris, C. A geographic information system (gis)-based analysis of social capital data: Landscape factors that correlate with trust. *Sustainability* **2017**, *9*, 365. [[CrossRef](#)]

15. Ottaviani, D.; Panzacchi, M.; Lasinio, G.J.; Genovesi, P.; Boitani, L. Modelling semi-aquatic vertebrates' distribution at the drainage basin scale: The case of the otter *lutra lutra* in Italy. *Ecol. Model.* **2009**, *220*, 111–121. [[CrossRef](#)]
16. Poplar-Jeffers, I.O.; Petty, J.T.; Anderson, J.T.; Kite, S.J.; Strager, M.P.; Fortney, R.H. Culvert replacement and stream habitat restoration: Implications from brook trout management in an appalachian watershed, USA. *Restor. Ecol.* **2009**, *17*, 404–413. [[CrossRef](#)]
17. Patel, R.P.; Wutke, S.; Lenz, D.; Mukherjee, S.; Ramakrishnan, U.; Veron, G.; Fickel, J.; Wilting, A.; Förster, D.W. Genetic structure and phylogeography of the leopard cat (*Prionailurus bengalensis*) inferred from mitochondrial genomes. *J. Hered.* **2017**, *108*, 349–360. [[CrossRef](#)] [[PubMed](#)]
18. Tan, S.; Xu, J.-T.; Zou, F.-D.; Peng, Q.-K.; Peng, R. The complete mitochondrial genome of leopard cat, *Prionailurus bengalensis chinensis* (carnivora: Felidae). *Mitochondrial DNA A* **2016**, *27*, 3073–3074.
19. Jang, K.H.; Hwang, U.W. Complete mitochondrial genome of korean yellow-throated marten, *Martes flavigula* (carnivora, mustelidae). *Mitochondrial DNA A* **2016**, *27*, 1785–1786. [[CrossRef](#)] [[PubMed](#)]
20. Chen, M.-T.; Liang, Y.-J.; Kuo, C.-C.; Pei, K.J.-C. Home ranges, movements and activity patterns of leopard cats (*Prionailurus bengalensis*) and threats to them in Taiwan. *Mamm. Study* **2016**, *41*, 77–86. [[CrossRef](#)]
21. Vigne, J.-D.; Evin, A.; Cucchi, T.; Dai, L.; Yu, C.; Hu, S.; Soulages, N.; Wang, W.; Sun, Z.; Gao, J. Earliest “domestic” cats in China identified as leopard cat (*Prionailurus bengalensis*). *PLoS ONE* **2016**, *11*, e0147295. [[CrossRef](#)] [[PubMed](#)]
22. Mohamed, A.; Ross, J.; Hearn, A.J.; Cheyne, S.M.; Alfred, R.; Bernard, H.; Boonratana, R.; Samejima, H.; Heydon, M.; Augeri, D.M.; et al. Predicted distribution of the leopard cat *Prionailurus bengalensis* (mammalia: Carnivora: Felidae) on borneo. *Raffles Bull. Zool.* **2016**, *33*, 180–185.
23. Lee, M.-J.; Song, W.; Lee, S. Habitat mapping of the leopard cat (*Prionailurus bengalensis*) in South Korea using gis. *Sustainability* **2015**, *7*, 4668–4688. [[CrossRef](#)]
24. Bui, D.T.; Lofman, O.; Revhaug, I.; Dick, O. Landslide susceptibility analysis in the hoa binh province of vietnam using statistical index and logistic regression. *Nat. Hazards* **2011**, *59*, 1413. [[CrossRef](#)]
25. Huck, M.; Davison, J.; Roper, T.J. Predicting European badger *meles meles* sett distribution in urban environments. *Wildl. Biol.* **2008**, *14*, 188–198. [[CrossRef](#)]
26. Gavashelishvili, A.; Lukarevskiy, V. Modelling the habitat requirements of leopard panthera pardus in west and Central Asia. *J. Appl. Ecol.* **2008**, *45*, 579–588. [[CrossRef](#)]
27. Northrup, J.; Stenhouse, G.; Boyce, M. Agricultural lands as ecological traps for grizzly bears. *Anim. Conserv.* **2012**, *15*, 369–377. [[CrossRef](#)]
28. Kuemmerle, T.; Perzanowski, K.; Chaskovskyy, O.; Ostapowicz, K.; Halada, L.; Bashta, A.-T.; Kruhlov, I.; Hostert, P.; Waller, D.M.; Radeloff, V.C. European bison habitat in the Carpathian mountains. *Biol. Conserv.* **2010**, *143*, 908–916. [[CrossRef](#)]
29. Clement, M.J.; Castleberry, S.B. Estimating density of a forest-dwelling bat: A predictive model for rafinesque's big-eared bat. *Popul. Ecol.* **2013**, *55*, 205–215. [[CrossRef](#)]
30. Greaves, G.J.; Mathieu, R.; Seddon, P.J. Predictive modelling and ground validation of the spatial distribution of the New Zealand long-tailed bat (*Chalinolobus tuberculatus*). *Biol. Conserv.* **2006**, *132*, 211–221. [[CrossRef](#)]
31. Maleki, S.; Soffianian, A.R.; Koupaei, S.S.; Saatchi, S.; Pourmanafi, S.; Sheikholeslam, F. Habitat mapping as a tool for water birds conservation planning in an arid zone wetland: The case study hamun wetland. *Ecol. Eng.* **2016**, *95*, 594–603. [[CrossRef](#)]
32. Novaczek, E.; Devillers, R.; Edinger, E.; Mello, L. High resolution habitat mapping to describe coastal denning habitat of a Canadian species at risk, atlantic wolffish (*Anarhichas lupus*). *Can. J. Fish. Aquat. Sci.* **2017**. [[CrossRef](#)]
33. Park, J.-K.; Das, A.; Park, J.-H. A new approach to estimate the spatial distribution of solar radiation using topographic factor and sunshine duration in South Korea. *Energy Convers. Manag.* **2015**, *101*, 30–39. [[CrossRef](#)]
34. Woo, D.G. A Study on Ecological Characteristics and Conservation of Yellow-Throated Marten (*Martes flavigula*) in Temperate Forests of Korea. Ph.D. Thesis, Seoul National University, Seoul, Korea, 2014.
35. Rumelhart, D.E.; McClelland, J.L.; Group, P.R. *Parallel Distributed Processing*; MIT Press: Cambridge, MA, USA, 1987; Volume 1.
36. Hines, J.; Tsoukalas, L.H.; Uhrig, R.E. *Matlab Supplement to Fuzzy and Neural Approaches in Engineering*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1997.

37. Althuwaynee, O.F.; Pradhan, B.; Lee, S. A novel integrated model for assessing landslide susceptibility mapping using chaid and ahp pair-wise comparison. *Int. J. Remote Sens.* **2016**, *37*, 1190–1209. [[CrossRef](#)]
38. Kim, J.-C.; Lee, S.; Jung, H.-S.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-chang, Korea. *Geocarto Int.* **2017**, 1–16. [[CrossRef](#)]
39. Lee, S.; Kim, J.-C.; Jung, H.-S.; Lee, M.J.; Lee, S. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomat. Nat. Hazards Risk* **2017**, 1–19. [[CrossRef](#)]
40. Tahmassebipoor, N.; Rahmati, O.; Noormohamadi, F.; Lee, S. Spatial analysis of groundwater potential using weights-of-evidence and evidential belief function models and remote sensing. *Arab. J. Geosci.* **2016**, *9*, 79. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).