Applicability of a predictive distributional model for the distribution of the Asiatic black bear (\textit{Ursus thibetanus japonicus}) to a local population in the Tanzawa region of Japan

ソキノワグマ（\textit{Ursus thibetanus japonicus}）生息予測分布モデルの丹沢地域個体群での適用可能性

Tomoko DOKO*, Wenbo CHEN** and Tomohiro ICHINOSE***


tokyo University

Abstract: A predictive distributional model of Asiatic black bears (\textit{Ursus thibetanus japonicus}) is useful for the effective management of the species’ habitat. Applicability of a previously developed model to the Tanzawa local bear population was examined. Three different types of animal presence data were collected for validation of the model. The overall rate of correct classification was 67\%–85\%, indicating that the model can be extrapolated to other local populations with acceptable transferability. Accuracy of the model’s predictions differed depending on the data sampling design most likely as a result of bias in one of the sampling methods. The model tended to underestimate the actual presence of bears. We attribute this underestimation to 1) properties of the particular habitat used by the local Tanzawa population and 2) the model’s greater ability to predict the habitat of mountain bears than that of Satoyama bears.

Keywords: Asiatic black bear, \textit{Ursus thibetanus japonicus}, applicability, predictive distributional model, Tanzawa local population

要旨: ソキノワグマ（\textit{Ursus thibetanus japonicus}）の生息予測分布モデルは有効な生息地管理に役立つ。本研究では、丹沢地域個体群に対するモデルの適用可能性を論じた。モデルの精度検証のために3種類の現地の在データを集めた。全体的なモデル精度は、正判率が67\%–85\%であり、これは、本モデルが許容できる転送性を持っており、他の地域個体群にも外挿できることを示唆している。サンプリング方法によって精度が異なったが、I種類のサンプリング方法によって生じたバイアスが原因であると解釈した。モデルには、実際にクマが生息する場所を過小評価する傾向があり、本地域個体群特有の生息地の特性による影響と、リグマと比較して山ダマへの高い予測能力があるためと思われた。

キーワード: ソキノワグマ、\textit{Ursus thibetanus japonicus}、適用可能性、予測分布モデル、丹沢地域個体群

INTRODUCTION

Estimating habitat use by the Asiatic black bear (\textit{Ursus thibetanus japonicus}, hereinafter “bear”) is crucially important for effective management because this species is considered to be both a prey animal and an umbrella species. Several previous studies (Doko et al., 2009; 2011) have attempted to develop predictive models to estimate the probability of bear occurrence using environmental indices, and these studies validated model accuracy using data on the presence and absence of the species. Doko et al. (2009) introduced logistic regression models to estimate the probability of occurrence of the bear, and Doko et al. (2011) developed a predictive distributional map using a Maximum Entropy (MaxEnt: Phillips et al., 2006) model for examining conservation gaps and estimating population size. MaxEnt requires only presence data (point localities in which a species is known to occur) and predictors, such as environmental parameters that are thought to limit or enhance the species’ chances of survival. The environmental parameters used in Doko et al. (2011) were partitioned into four categories: topology, water resources, distance to roads, and vegetation.

The transferability of a model reflects its ability to predict species distribution in broad, unsampled regions (Peterson et al., 2007). Transferability tests are important for several types of applications, including predicting the effects of global climate change on species potential distribution, evaluating the invasive potential of an alien species, and the discovery of new populations and species (Peterson et al., 2007).

In Honshu and Shikoku, Japan, there are 19 recognized local bear populations (Yui et al., 1994). The transferability of a distributional model, initially developed using data on a local population in the South Alps of central Japan was tested using the neighboring local population in Fuji and high transferability was reported (Doko et al., 2009; 2011). Transferability test of the model for another local population is a useful next step.
This study aims to test the transferability of the predictive distributional model of the bear (hereinafter, "model") using another local population, the Tanzawa population. The model was developed in Doko et al. (2011) using MaxEnt. We refer the reader to that paper for a description of the development of the model. The model was trained using presence data from 14 tracked individual bears in combination with environmental parameters such as altitude, distance to paths and stone steps, distance to wide roads (>13 m in width), and vegetation cover types (at the community level) (Doko et al., 2011). In this study, the model was tested using field data collected in the Tanzawa region. Finally, the applicability and limitations of the model are discussed.

1. METHODS

1.1 Study area

The study area was the Tanzawa region located in the northwestern area of the Kanagawa Prefecture in eastern Japan (35.20°-35.49°N and 138.55°-139.20°E, Fig. 1A). The Tanzawa region is a mountainous area with the highest mountain, Mt. Hirugatake, at 1,673 m and 9 other mountains with altitudes that exceed 1,500 m, including Mt. Tanzawa and Mt. Hinokiboramatsu (Research Group of the Tanzawa Mountains 2007). The land cover of this region is mainly forest; the Tanzawa mountains have an area of 40,000 ha, which accounts for up to 16% of the total area of the Kanagawa Prefecture (Research Group of the Tanzawa Mountains 2007). This area is designated as Tanzawa Quasi-National Park. The vegetation is characterized by virgin forest of Japanese Beech (Fagus crenata) and Japanese Fir (Abies firma). The Tanzawa Mountains at lower elevations have a high percentage of planted forests such as Japanese Cedar (Cryptomeria japonica) and Hinoki cypress (Chamaecyparis obtusa) (Research Group of the Tanzawa Mountains, 2007). The local bear population in the Tanzawa region is estimated to be 26 to 30 animals (Hazumi et al., 1997; Doko et al., 2011); hence it is critically endangered.

1.2 Preparation of presence data for validation

1) Body hair data

The Kanagawa Prefectural Natural Environment Conservation Center outsourced hair trap surveys of bears in the Tanzawa Mountains to the Wildlife Management Office Inc. (Tokyo, Japan). Thirty-five individual bears were identified by DNA analysis in the three years between 2006 and 2008. From internal reports, 49 locality points of bear body hair and other field signs, which are considered to be evidence of presence of bears, were collected (the number of hair traps was 20, 63, and 66 in 2006, 2007,
and 2008, respectively).

2) Location data from a collared bear: B
From 2008 to 2010, three traps in the form of double-barreled drum cans were set in the forest surrounding Mt. Hemuro near Miyagase Lake in Kiyokawa village, Kanagawa Prefecture. One-half liter of honey was placed in each trap as bait. We captured an adult female bear (named HANA, microchip ID: 392141000116744, estimated to be 4 years old, 43 kg in weight, and 123 cm in body length) on 13 September 2009 and fitted her with a GPS-ARGOS collar (ARGOS97801, model TGW-4580, Telonics Inc.). From 19 September to 23 October 2009 (35 days, autumn season), this bear's location was regularly transmitted to the ARGOS server, and 106 separate GPS location points were collected for the bear through the server (Doko et al., unpublished).

3) Field signs: C
From 2007 to 2009, we collected data on field signs of bears in Kiyokawa village in the Tanzawa Mountains, based on "sign surveys". Although it is best to design sign surveys by setting up fixed transects, for threatened or rare mammals this is not feasible. This survey considered sample data from a variety of environments such as broad leafed forests and plantations. Field signs for which data were collected included scats, footprints, bear nests, scratching posts, bear bark stripping, and body hair. The location of each field sign was recorded using a portable GPS receiver (eTrex LEGEND HCx, GARMIN), and photographs were taken of each field sign. Bear nests are constructed by the bears in trees that they climb to eat fruits and nuts (Maita, 1996). Bear bark stripping is defined as the bears' behavior of peeling the bark off of evergreen trees (Maita, 1996). Scratching posts are sections of trees that the bears climb up and down, scratching the wood with their claws (Maita, 1996). The locations where bear body hair was found were recorded when the authors participated in the survey of the Kanagawa Prefecture. In addition, two local hunters interviewed in 2007 and in 2008 provided the locations of bear hibernation sites and habitats in which bears were frequently observed in Mt. Hemuro (Fig. 1B) and in Yamakita village (Fig. 1A). In total, 25 field signs of bears were collected to serve as presence data for the species in the Tanzawa area.

1.3 Preparation of absence data
Test data regarding the absence of the species were prepared from Doko et al. (2011), using only the data from that study for the Tanzawa region. The total sample size for test data regarding absence was 95.

1.4 Validation of accuracy
The validation method used here is the same as that in Doko et al. (2011). Accuracy was measured using Kappa statistics and AUC (area under the curve). The cut-off probability (hereafter, "optimum probability"), which can maximize the correct classification rate, was calculated with the program "ROC Plotting and AUC Calculation Transferability Test Version 1.3.7" ①. Using R statistical software version 2.9.1 (R Development Core Team, 2005), indices including Kappa, correct classification rate, bias index, prevalence index, sensitivity, specificity, and AUC were calculated at the optimum probability.

2. RESULTS

2.1 Geographic location of presence/absence data
Fig. 1A shows a map of presence data points with the predicted presence area for bears in the local Tanzawa region population. Fig. 1B shows where the predictions failed around Mt. Hemuro in Kiyokawa village of the Tanzawa Mountains. As seen in Fig. 1A, most of the field signs of body hair (A) and the telemetry data from the tracked bear (B) were located inside of the predicted presence area. However, on the eastern side of Tanzawa, in the vicinity of Mt. Hemuro, the locations of field signs (C) fell outside of the predicted presence area (Fig. 1). Field signs spatially overlapped with habitats in which bears had been observed by a hunter, near the center of Kiyokawa village (1-2 km) at a lower altitude (400-500 m; Fig. 1B). Based on the hair trap sampling localities (A) and telemetry data from the tracked bear (B), Mt. Hemuro lies at the eastern limit of the local Tanzawa population distribution (see Fig. 1A).
2.2 Model validation

The confusion matrix is shown in Table 1 for field signs collected by the Kanagawa Prefecture (A), location data from the tracked bear (B), and field signs collected by the authors (C). The accuracy of the model based on each type of presence data was compared with the accuracies obtained in previous studies by Doko et al. (2011) and Doko et al. (2009) (Table 2). Among these three types of data, the field signs collected by Kanagawa Prefecture (A) achieved the highest accuracy (Kappa K=0.65, Correct classification rate (CCR)=0.85, Tables 1 and 2). The tracked bear’s location data (B) also yielded high accuracy (K=0.53, CCR=0.76, Tables 1 and 2). In contrast, the presence data based on field signs collected by the authors (C) decreased the total accuracy; and all of these data (25 points) were incorrectly classified by the model (Table 1). Therefore, accuracy of the field sampling design was calculated without these data for reference (Tables 1 and 2, A+B). For “A+B” sampling design, CCR, K, and AUC were 0.74, 0.50, and 0.86, respectively.

The optimum probability was p=0.005 for sampling designs of “A” and “B”, and p=0.525 for sampling design “C”. For the combined data (A+B+C in Table 1), CCR was 0.67, and K was 0.39. Of a total of 180 actual presence data points, 86 were incorrectly classified as absences. Absence data were classified with less error: 4 points of a total of 95 actual absence data points were incorrectly classified as presences.

3. DISCUSSION

In the original MaxEnt model, the optimal probability was p=0.005 for test data from the Fuji local population (Table 2, Doko et al., 2011). In this study, the optimal probability was also p=0.005 for data collected with the total sampling design (A+B+C). These results are in agreement with Peterson et al. (2007) who reported that the MaxEnt models fails to make general predictions unless very low probability value thresholds are considered. In contrast, the logistic regression model (Doko et al., 2009) yielded higher threshold values of p=0.613 and p=0.033 for training data in the South Alps and test data in Fuji (Table 2). The indices regarding accuracy of training data in the MaxEnt and logistic regression models show almost perfect scores, but when the test data are compared to the training data, the accuracy apparently decreases in both models (Table 2). This indicates that these models probably over-fit the

<table>
<thead>
<tr>
<th>Table 1 Confusion matrix</th>
<th>A</th>
<th>At a threshold of 0.005</th>
<th>Predicted (+)</th>
<th>Actual (+)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted (-)</td>
<td>32</td>
<td>4</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-)</td>
<td>17</td>
<td>91</td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>47</td>
<td>95</td>
<td>144</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>At a threshold of 0.525</th>
<th>Predicted (+)</th>
<th>Actual (+)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted (-)</td>
<td>62</td>
<td>4</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>(-)</td>
<td>44</td>
<td>91</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>106</td>
<td>95</td>
<td>201</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>At a threshold of 0.525</th>
<th>Predicted (+)</th>
<th>Actual (+)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted (-)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(-)</td>
<td>25</td>
<td>95</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>25</td>
<td>95</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A+B</th>
<th>At a threshold of 0.005</th>
<th>Predicted (+)</th>
<th>Actual (+)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted (-)</td>
<td>94</td>
<td>4</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>(-)</td>
<td>61</td>
<td>91</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>155</td>
<td>95</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A+B+C</th>
<th>At a threshold of 0.005</th>
<th>Predicted (+)</th>
<th>Actual (+)</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted (-)</td>
<td>94</td>
<td>4</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>(-)</td>
<td>86</td>
<td>91</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>180</td>
<td>95</td>
<td>275</td>
<td></td>
</tr>
</tbody>
</table>

Note, A: Field signs collected by Kanagawa Prefecture, B: location data by a tracked bear by authors, and C: field signs collected by authors. (+) and (-) denote ‘presence’ and ‘absence’ respectively.

<table>
<thead>
<tr>
<th>Table 2 Accuracy by Kappa statistics and AUC at the optimized probability as a threshold</th>
<th>Optimized p</th>
<th>K</th>
<th>CCR</th>
<th>BI</th>
<th>PI</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimized p</td>
<td>0.005</td>
<td>0.65</td>
<td>0.85</td>
<td>0.090</td>
<td>0.409</td>
<td>0.6531</td>
<td>0.9679</td>
<td>0.87</td>
</tr>
<tr>
<td>B</td>
<td>0.005</td>
<td>0.53</td>
<td>0.76</td>
<td>0.199</td>
<td>0.144</td>
<td>0.5849</td>
<td>0.9679</td>
<td>0.85</td>
</tr>
<tr>
<td>C</td>
<td>0.525</td>
<td>0.00</td>
<td>0.79</td>
<td>0.208</td>
<td>0.791</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.53</td>
</tr>
<tr>
<td>A+B</td>
<td>0.005</td>
<td>0.50</td>
<td>0.74</td>
<td>0.226</td>
<td>0.012</td>
<td>0.6065</td>
<td>0.9679</td>
<td>0.86</td>
</tr>
<tr>
<td>A+B+C</td>
<td>0.005</td>
<td>0.39</td>
<td>0.67</td>
<td>0.298</td>
<td>0.010</td>
<td>0.5222</td>
<td>0.9679</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Note, K=Kappa, CCR=Correct classification rate, BI-Bias index, PI-Prevalence index. A: Field signs collected by Kanagawa Prefecture, B: location data by a tracked bear by authors, and C: field signs collected by authors. Test Maxent and train Maxent denote test data in Fuji and train data in South Alps, respectively (MaxEnt: Doko et al., 2011), while test logistic and train logistic denote test data in Fuji and train data in South Alps, respectively (logistic regression model) (Doko et al., 2009).
training data. In this study, the overall correct classification rate was 67% (CCR=0.67, Table 2). When sampling design “C” was excluded from the analysis, the CCR became 74% (Table 2). However, Kappa (κ) values were relatively low in both sampling designs (A+B+C: κ=0.39, A+B: κ=0.50, Table 2). Compared to the high applicability of the model to the Fuji local population (κ=0.75 and CCR=0.88, Table 2, Doko et al. (2011)), the accuracy obtained in the present study was lower. The lower Kappa values are thought to be caused by a high level of omission error. Nevertheless, transferability of the model to the Tanzawa local population is acceptable (~67–85% correctness, Table 2). In particular, if the model predicts the presence of bears, bears most likely are present, because few commission errors occurred. In general, we have found that the model can be extrapolated to other local populations with acceptable transferability. However, due to differences in habitat conditions, extrapolation should be done carefully. For instance, it is possible that heavy snow region influences the bears’ habitat selection.

Accuracy of the model predictions differed depending on the sampling designs used to obtain test data. Accuracy was lowest for the field signs collected by the authors (C), whereas hair trap survey data (A) and telemetry data (B) yielded relatively high accuracy (Table 2). This difference is thought to be caused by sampling bias in “C”. We assume the surveyors of hair traps (A) knew the habitats and the range of bears in the Tanzawa region because of previous experiences with telemetry in this region (the same research team was used by Hazumi et al., 1997). Based on this prior knowledge, the surveyors divided the area into regional sections (Kiyokawa village, Kyu-Tsukui village, Hadano city, Matsuda village, and Yamakita village) and conducted a survey across the entire Tanzawa region. The GPS-ARGOS telemetry method used in this study regularly and automatically collected location data from a collared bear, and we believe this method successfully revealed the whole range of one bear without bias. In contrast, the authors collected field signs in a biased manner. First, we surveyed only one region in Kiyokawa village, which was found to be located on the eastern fringe of the Tanzawa local population. Second, because a vehicle was used to transport heavy traps, all of the selected trap locations were close to roads. Finally, our survey to find field signs of bears was carried out near the traps, as a supplemental investigation to the telemetry survey.

For all field sampling strategies (A, B, and C), the same pattern of error in the model’s predictions is observed. This is characterized by little commission error, or few “false positives” and much omission error, or many “false negatives”. Thus, the model’s predictions tend to underestimate the actual presence of bears and incorrectly classify data points as absences in the Tanzawa local population (Table 1). Why does this underestimation occur?

Firstly, we attribute this omission error to a particular feature of the habitat used by the Tanzawa local population. The failed predictions are concentrated in Kiyokawa village, where the field signs collected by the authors ((C) in Fig. 1) were located close to the roads in Mt. Hemuro. Through analysis of the South Alps local population, Doko et al. (2009) found that bears avoid roads. In fact, in the Tanzawa region, there are many roads with restricted access that are used only by permitted forestry personnel (check marks, Fig. 1B). This probably makes the volume of traffic on these roads very low, and bears appear not to avoid these roads. Nevertheless, in the model, the two predictors 1) proximity to paths and stone steps and 2) proximity to wide roads (>13 m in width) were negatively associated with bear habitat selection (Doko et al., 2009, 2011). This suggests that the places in the Tanzawa region where the model predicts low probability of bear occurrence due to road proximity are in fact bear habitats, due to the positive influence of the traffic restriction on these roads.

Secondly, our results suggest that the model more accurately predicts the habitats of mountain bears than that of Satoyama bears. Two types of bears with distinct behavioral patterns are thought to exist in Japan: mountain bears, which inhabit remote mountains, and Satoyama bears, which inhabit areas close to human settlements (Maita, 1998). Based on field signs collected by the authors, bears certainly utilize resources near the village. The telemetry data from a collared bear (HANA) indicate that HANA's
range included both Satoyama and mountainous regions (Fig. 1). One possibility is that the localities of mountain bears (on the west side of the Tanzawa Mountains, Fig. 1A) are correctly classified because altitude was used in the model as a positive predictor of bear survival. However, in Kiyokawa village, bears were observed near human settlements (Fig. 1B). This indicates that the bears also utilize resources located closer to human settlements and at lower altitudes, as well as resources in the predicted habitats distant from settlements.

CONCLUSION

Underestimation of bear ranges can lead to conflicts between humans and bears with respect to space use because bears are considered pest animals. Kiyokawa village is surrounded by mountainous forests, and human settlements are concentrated in small, flat areas (see Center of Kiyokawa village in Fig. 1B). The designation of buffer zones is therefore recommended, preferably around Mt. Hemuro, between Kiyokawa village and the bears’ habitat areas. Future development of an additional model that correctly predicts habitats of Satoyama bears would enable us to create hazard maps as well as to quantify the habitat conditions preferred by these Satoyama bears.

The underestimation that occurred in this study also indicates that the number of bears in the Tanzawa local population is larger than previously estimated. Doko et al. (2011) estimated that 26 bears inhabit the Tanzawa region based on the size of the habitat area and population density. Because the present study reveals that the potential habitat extends east to Mt. Hemuro, we now believe that the number of bears in the Tanzawa local population is more than 26.

ACKNOWLEDGEMENTS

This research was financed by a Project Grant-in-Aid from the SFC Research Institute in 2007, Keio University, the Research promotion fund in remembrance of Mori Takichiro (2007/2008), the Graduate School Doctoral Students Grant-in-Aid Program (2008), and a Grant-in-Aid for JSPS Fellows (2009, KAKENHI No.215601). The authors thank Mr. Koji Obara, Mr. Hajime Sugimoto, and Mr. Yasuharu Watabe (Hunting Association), Mr. Hideo Higushi, the Staff of Kiyokawa village/Kanagawa Prefecture, Ms. Kayoko Suetugu (KPNCC), Assoc. Prof. Satoshi Osawa (Nihon University), and Mr. Masayuki Saito (Yokohama National University).

NOTES


REFERENCES


