Habitat of the Vulnerable Formosan sambar deer *Rusa unicolor swinhoii* in Taiwan

Shih-Ching Yen, Ying Wang and Heng-You Ou

Abstract The sambar deer *Rusa unicolor* is categorized as Vulnerable on the IUCN Red List because of continuous population decline across its native range. In Taiwan the Formosan sambar deer *R. unicolor swinhoii* is listed as a protected species under the Wildlife Conservation Act because of human overexploitation. However, its population status remains unclear. We used presence and absence data from line transect and camera-trap surveys to identify key habitat variables and to map potential habitats available to this subspecies in Taiwan. We applied five habitat-suitability models: logistic regression, discriminant analysis, ecological-niche factor analysis, genetic algorithm for rule-set production, and maximum entropy. We then combined the results of all five models into an ensemble-forecasting model to facilitate a more robust prediction. This model indicated the existence of 7,865 km^2 of suitable habitat for the sambar deer. Distance from roads and elevation were identified as the most important environmental variables for habitat suitability, and deer preferred areas far from roads and $>$ 1,500 m altitude. The results predicted that suitable deer habitat is mainly located in Taiwan's Central Mountain Range and Xue Mountain Range, with c. 70% of this suitable habitat in protected areas. However, the habitat predicted to be suitable is in five areas separated by mountain highways. We recommend that deer habitats close to the highways should be monitored for the future establishment of corridors between Formosan sambar deer sub-populations.

Keywords Habitat suitability, predicted distribution, *Rusa unicolor*, sambar deer

Introduction

Large wild herbivores are primary consumers that play an important role in ecosystems and provide a substantial economic resource for many communities. However, human land use has caused fragmentation, degradation and loss of habitat for these species (Ceballos & Ehrlich, 2002). Furthermore, as a result of more effective hunting techniques, overexploitation has become the most important threat after habitat destruction for the survival of large herbivores (Groom, 2006). Consequently many ungulates in Asia are confined to protected areas and are limited to small populations (Baskin & Danell, 2003). Thus conservation actions are required to ensure the long-term survival of these animals.

One of the most important factors in successfully managing and conserving a species is accurate identification of its distribution (Boitani et al., 2008). To accomplish this we must understand how environmental factors determine the habitat of a species. Advances in habitat suitability modelling techniques combined with geographical information systems provide accessible tools for identifying suitable habitats and predicting potential species distribution (Anderson et al., 2003; Gavashelishvili & Javakhishvili, 2010). However, predictions of species distribution can vary significantly with different modelling techniques (Thuiller et al., 2004; Pearson et al., 2006) and it is difficult to select the most realistic. Araújo & New (2007) advocated the use of ensemble forecasting, which often generates a more robust prediction than a single model (Araújo & New, 2007; Marmion et al., 2009; Thuiller et al., 2009; Oppel et al., 2012).

The sambar deer *Rusa unicolor* is a large ungulate distributed throughout South and South-East Asia (Leslie, 2011). Although this species became a pest after introduction to countries such as Australia (Gormley et al., 2011), populations in its native range have declined, with many local-level extinctions as a result of extensive hunting and habitat loss (Timmins et al., 2008). It remains regionally abundant only in well-secured (i.e. protected or remote) areas. The sambar deer is categorized as Vulnerable on the IUCN Red List (Timmins et al., 2008). Although some countries, such as Thailand, have banned hunting, the recovery rate of sambar deer populations remains slow, requiring further evaluation of population distributions and dynamics (Steinmetz et al., 2009).

The Formosan sambar deer *R. unicolor swinhoii* is a subspecies endemic to Taiwan (Wilson & Reeder, 2005). It is categorized as a rare and valuable species in the List of Protected Species in Taiwan (Forest Bureau of Taiwan, 2009). In the decades before the 1990s there was a decline in the number and geographical distribution of the Formosan sambar deer in Taiwan, reflecting similar trends in other areas of the sambar deer's range. However, in the mid 1990s there was a slight increase in Formosan sambar deer

Shih-Ching Yen and Ying Wang Department of Life Science, National Taiwan Normal University, Taipei, Taiwan

Heng-You Ou (Corresponding author) Institute of Ecology and Evolutionary Biology, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan. E-mail pinpon1216@gmail.com

Received 23 April 2012. Revision requested 10 October 2012.
Accepted 10 October 2012. First published online 17 October 2013.

© 2013 Fauna & Flora International, *Oryx*, 48(2), 232–240 doi:10.1017/S0030605312001378

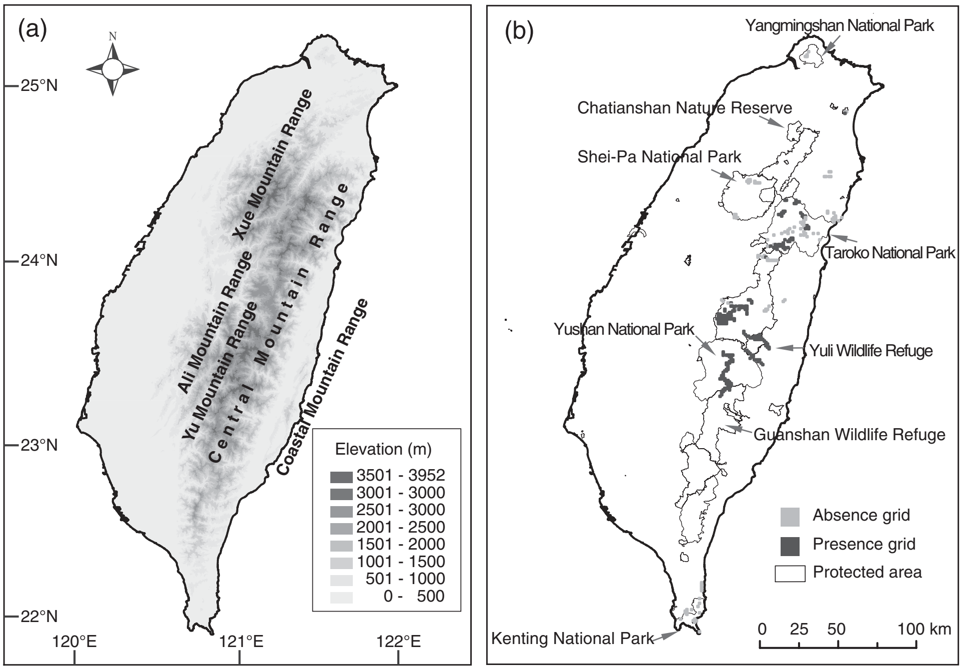

FIG. 1 Map of Taiwan, showing (a) the locations and elevations of the mountain ranges and (b) the locations of protected areas and recorded locations of Formosan sambar deer in Taiwan. Location data were collected from field surveys (2008–2012) and assimilated from previous studies (2002–2007).

populations in Taiwan (Timmins et al., 2008). Surveys indicated that Formosan sambar deer were mainly distributed in the Central Mountain Range (Wang et al., 2002; Fig. 1a), although the surveys were limited in effort relative to the size of the island.

Several studies in India (Porwal et al., 1996; Kushwaha et al., 2004) and Australia (Forsyth et al., 2009; Gormley et al., 2011) have described the distribution patterns and habitat selection of sambar deer. In Taiwan suitable habitat for the species lies above 1,000 m (Wang et al., 2002). In addition, the climate and vegetation types are subject to variation along the elevation gradient and are different to those of other countries. Furthermore, Taiwan has high population and road densities, which have caused extensive habitat destruction. Therefore investigations of habitat selection by sambar deer in highly disturbed areas are required. Our aims in the study reported here were to evaluate habitat suitability and to map the distribution pattern of Formosan sambar deer throughout Taiwan. We use the results to identify the most important sites for the management of this threatened subspecies.

Study area

The 35,801 km² island of Taiwan lies off the south-eastern coast of mainland China. Topography is high and steep, with five mountain ranges (Fig. 1a), and altitudes of 0–3,952 m. The vegetation changes with altitude, from broadleaf forests to coniferous forests and then to scrub (Su, 1992). Two-thirds of the island is covered by forested mountains. Most of the coastal plains are occupied by human settlements. There are 89 protected areas, covering a total of c. 6,951 km² (Fig. 1b). The climate is tropical marine, with warm and humid weather (mean annual temperature in the lowlands is c. 23 °C; mean annual precipitation is > 2,500 mm; Central Weather Bureau of Taiwan, 2013). However, the weather is cold in the high mountains, where it snows during winter. Snow cover duration is short, from several days to 2 months, depending on height and latitude.

Methods

Data collection

The locational data were primarily assimilated from our field observations of sambar deer (2008–2012). Data from field studies at other sites were also included (Pei et al., 2002, 2003; Pei, 2004; Wu & Shi, 2006; Lee et al., 2007; Wu & Yao, 2007; C.Y. Lin, pers. comm.). Although the collection of multiple datasets using different techniques prevented a standardized evaluation, it allowed the incorporation of data from many sites and environments. Two different survey techniques were used. The first involved the use of line transects to obtain data on deer presence. Absence data were not collected because it is difficult to confirm the absence of a species in such surveys. Transect lines were located on hiking and hunting trails. At each study site we surveyed several transect lines, to cover the various environments of that site. We recorded 1,582 coordinates of sambar deer tracks and signs (i.e. sightings, vocalizations, scats, footprints, tree rubbing and shed antlers) using a global positioning system. The second technique involved the use

Table 1 Environmental variables used to predict the distribution of Formosan sambar deer *Rusa unicolor swinhoii* in Taiwan, their values, and the source of the data.

Variable	Value	Source
Mean elevation	0–3,706 m	Ministry of the Interior, Taiwan
Standard deviation of elevation	0–350.19	Ministry of the Interior, Taiwan
Distance to water body	0–8 km	Institute of Transportation, Taiwan (2008)
Annual mean temperature	6.5–25.17 °C	Central Weather Bureau, Taiwan (1990)
Annual precipitation	1,179–5,700 mm	Central Weather Bureau, Taiwan (1990)
Vegetation type	7 classes	1:50,000 Editorial committee of the Flora of Taiwan (2nd edition 1994)
Forest area	0–1 km^2	1:50,000 Editorial committee of the Flora of Taiwan (2nd edition 1994)
Road density	0–66.60 km/km^2	Institute of Transportation, Taiwan (2008)
Distance to road	0–22.09 km	Institute of Transportation, Taiwan (2008)
Human settlement cover	0–1 km^2	Forestry Bureau, Taiwan (1995)

of 258 camera traps to obtain deer presence–absence data. Camera traps were laid 10–100 m away from the transect lines on animal trails. Camera-trap photographs of sambar deer were classified as presence data. If a camera trap operated for > 20 days without any photograph of sambar deer, this was regarded as a record of absence. In total 1,840 records of sambar deer were gathered, comprising 1,645 presence records and 195 absence records.

These records were transformed to a resolution of 1 km^2; i.e. records within the same grid were incorporated into a single presence or absence record. A total of 361 grid cells of 1 km^2 were sampled, representing 1% of the total land area of Taiwan. Line transect records that overlapped with grid cells categorized as 'absent' were re-examined, and five of these grid cells were recategorized as 'presence'. Overall, 241 of the grid cells had records of presence and 120 of absence (Fig. 1b).

Environmental variables

We used 10 environmental variables with potential importance for sambar deer habitat suitability (Kushwaha et al., 2004; Forsyth et al., 2009; Gormley et al., 2011): mean elevation, standard deviation of elevation, distance to water body, annual mean temperature, annual precipitation, vegetation type, forest area, road density, distance to road, and human settlement cover (Table 1). All variables were obtained from the ecological and environmental geographical information system database for Taiwan (Lee et al., 1997) and transformed to a resolution of 1 × 1 km using *ArcGIS 9.3* (ESRI, Redlands, USA) and *IDRISI Andes* (Clark Labs, Worcester, USA).

Statistical analyses and development of models

We divided the data into training and testing data. We used the heuristic method provided by Huberty (1994) to determine the ratio of testing data to the complete data set:

$$1/[1+\sqrt{(p-1)}]$$

where p is the number of environmental variables. We used 10 environmental variables, therefore the ratio of training to testing data should be 3:1. We analysed the distribution of the sambar deer using logistic regression, discriminant analysis, ecological-niche factor analysis, genetic algorithm for rule-set production, and maximum entropy (see Appendix for descriptions of each). Logistic regression and discriminant analysis are presence–absence models, and the other three are presence-only models. These models are regularly used to predict species distributions (Teixeira et al., 2001; Hirzel et al., 2002; Brotons et al., 2004; Lee et al., 2006).

We used different methods for each model to determine the importance of each environmental variable for the distribution of sambar deer. In the analysis of maximum likelihood estimates we used Wald χ^2 statistics; in discriminant analysis we used standardized canonical discriminant function coefficients; for ecological-niche factor analysis we used the factor scores; and for the maximum entropy model we used (1) jack-knife analysis of the mean gain with the training and test data, in addition to the area under the receiver-operating-characteristic curve (AUC), and (2) the mean percentage contribution of each environmental variable (Phillips et al., 2006). We were unable to evaluate the importance of variables for the genetic algorithm for rule-set production because of software limitations.

To evaluate the performance of each model, we used the AUC (Fielding & Bell, 1997), plotting the true-positive fraction against the false-positive fraction for all test points across all possible probability thresholds. The area under the curve measurement takes values between 0 and 1, with a value of 0.5 indicating that a model is no better than random. It is independent of prevalence and is considered an effective measure of the performance of ordinal score models (Manel et al., 2001; McPherson et al., 2004).

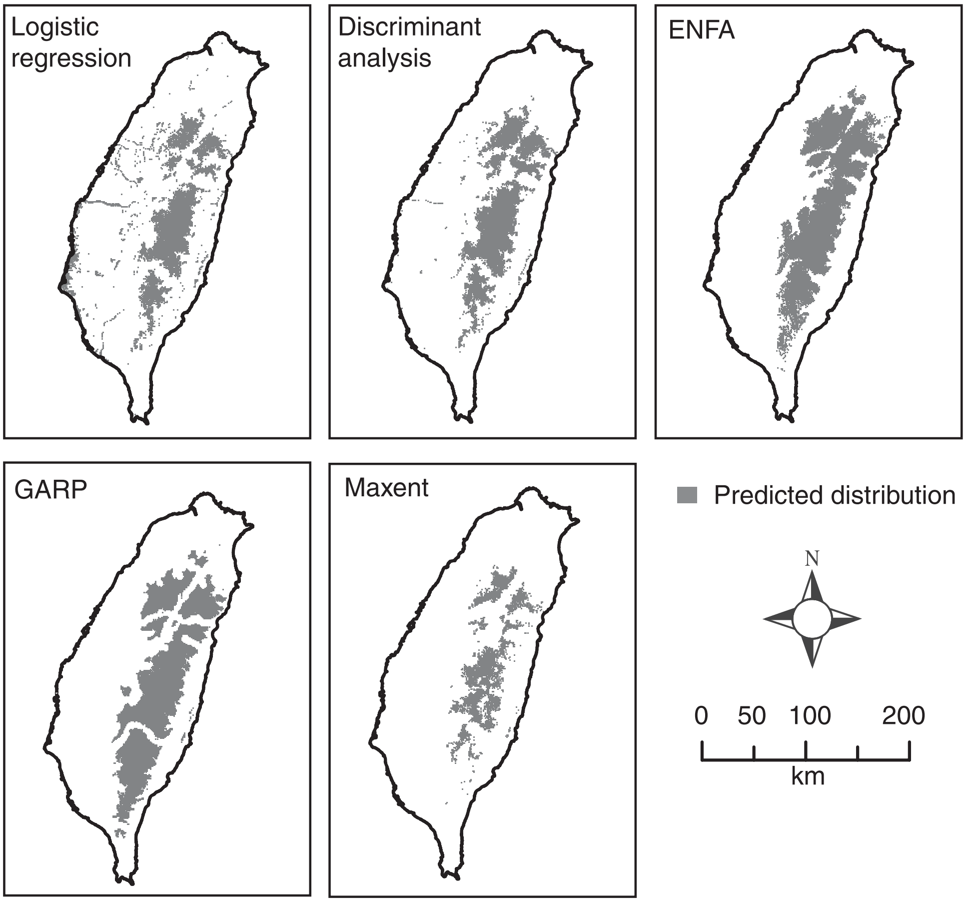

FIG. 2 Predicted habitat of Formosan sambar deer *Rusa unicolor swinhoii* in Taiwan, using (a) logistic regression, (b) discriminant analysis, (c) ecological-niche factor analysis, (d) genetic algorithm for rule-set production and (e) maximum entropy.

For conservation purposes it is usually desirable to distinguish suitable from unsuitable areas by setting a threshold. If the predicted probability of occurrence is larger than the threshold then it is considered to be a prediction of presence (Pearson et al., 2004). We calculated kappa statistics under different probabilities of occurrence and selected the probability that generated the maximum kappa statistic as the threshold for each model (Freeman & Moisen, 2008).

To obtain the most robust prediction map we used ensemble forecasting, as described by Araújo & New (2007). We calculated ensemble forecasting as a weighted mean by weighting each model based on its area under the curve (Araújo & New, 2007; Marmion et al., 2009; Thuiller et al., 2009; Oppel et al., 2012). The habitat suitability indices of ensemble forecasting ranged from 0 to 1. We summarized the areas of suitable habitat and optimal habitat using the suitability index thresholds 0.33 and 0.67, respectively. We selected these thresholds based on our knowledge of the Formosan sambar deer in Taiwan.

Results

The five habitat suitability models had areas under the curve of 0.894, 0.885, 0.807, 0.777 and 0.908. Each model predicted a different distribution pattern for the sambar deer (Fig. 2) but all except the genetic algorithm for rule-set production indicated that distance to road and the mean elevation are the most important factors predicting habitat suitability for sambar deer (Table 2). A composite map was produced by ensemble forecasting (Fig. 3). The results showed that ensemble forecasting performed better than any of the individual models (AUC = 0.921).

There were 7,865 grid cells categorized as suitable habitat for the sambar deer, of which 4,464 were regarded as optimal habitat. The most suitable deer habitat is in the Central Mountain Range and Xue Mountain Range; c. 70% (5,355 of 7,865 km^2) of suitable habitat lies in protected areas. The ensemble model indicated that sambar deer prefer habitat at medium to high elevation (> 1,500 m) and areas that lie away from roads. The mean elevation of suitable habitat is 2000 ± 600 m, with the predicted distribution including all areas > 3,000 m. The mean distance of suitable habitats to roads is 8.5 ± 4.6 km. In general, the suitability of habitat for sambar deer increases with increasing elevation and distance from roads.

There are five main patches of suitable habitat for sambar deer in Taiwan (Fig. 3). Two of these patches are in the Xue Mountain Range and the others are in the Central Mountain Range and Yu Mountain Range. These patches are separated by three major highways: the Central Cross-Island Highway,

TABLE 2. Ranks of contributions of environmental variables in four habitat suitability models: logistic regression, discriminant analysis, ecological-niche factor analysis and maximum entropy. The genetic algorithm for rule-set production could not be used to compare the gain contributions of each variable. Ecological-niche factor analysis could not be used to compute nominal variables, therefore the variable 'vegetation type' was excluded from this model.

Rank of contributions	Logistic regression	Discriminant analysis	Ecological-niche factor analysis	Maximum entropy
1	Distance to road	Distance to road	Mean elevation	Mean elevation
2	Mean elevation	Mean elevation	Distance to road	Distance to road
3	Annual precipitation	Annual mean temperature	Annual mean temperature	Vegetation type
4	Human settlement cover	Vegetation type	Standard deviation of elevation	Annual mean temperature
5	Distance to water body	Human settlement cover	Human settlement cover	Annual precipitation
6	Forest area	Road density	Road density	Forest area
7	Annual mean temperature	Distance to water body	Forest area	Road density
8	Road density	Forest area	Annual precipitation	Human settlement cover
9	Standard deviation of elevation	Standard deviation of elevation	Distance to water body	Standard deviation of elevation
10	Vegetation type	Annual precipitation		Distance to water body

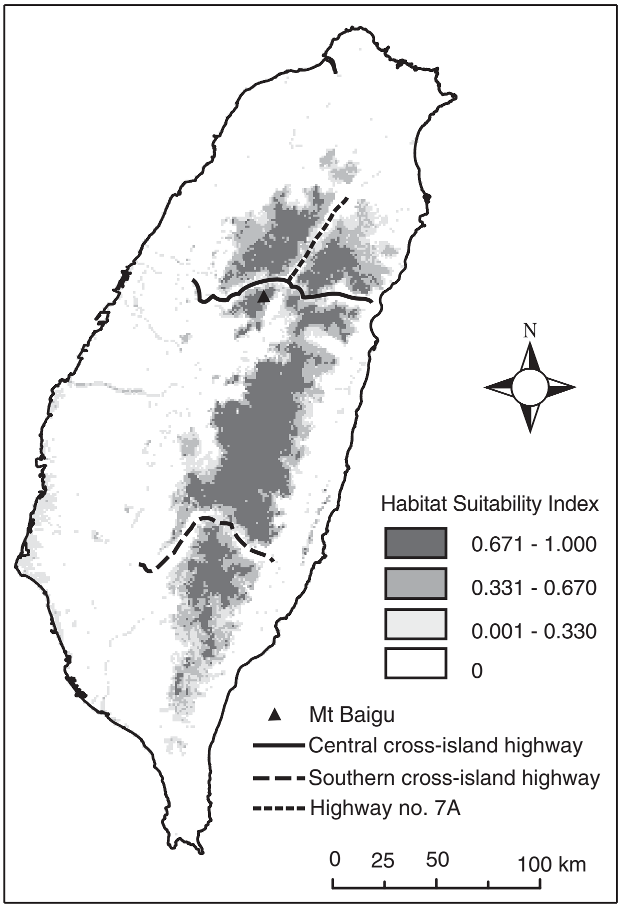

FIG. 3 Predicted habitats of Formosan sambar deer *R. unicolor swinhoii* based on ensemble forecasting, with three highways crossing the habitats.

the Southern Cross-Island Highway and Highway No. 7A (Fig. 3). In addition, three small patches of suitable habitat are located in the Ali Mountain Range, the Coastal Mountain Range and the Chatianshan Nature Reserve (Figs 1 & 3). Suitable habitat in areas of low elevation is scarce.

Discussion

The sambar deer has not been recorded in all of the 7,865 km^2 of habitat in Taiwan predicted to be suitable. For example, deer have not been detected in the Ali Mountain Range (Lin, 1997) or Chatianshan Nature Reserve (Wang, 1994). There are large areas of suitable habitat in the Xue Mountain Range but the population of sambar deer there is small (Fig. 1b; Yen, unpubl. data). Absence or sparse occurrence of the sambar deer is probably a result of high hunting pressure. An aboriginal tribe at Chatianshan and in the Xue Mountain Range has a long and prevalent hunting tradition, and local wildlife resources are often overexploited. Reintroduction (e.g. Klar et al., 2008; Kuemmerle et al., 2010) of the sambar deer to suitable unoccupied habitats would not necessarily be appropriate because the species is not under immediate threat of extinction in Taiwan, and hasty introductions could cause unanticipated damage to the local environment (Côté et al., 2004). We believe that monitoring the expansion of sambar deer populations and any associated environmental impacts is a more appropriate management technique at present. Of the 7,865 km^2 of habitat predicted to be suitable, 30% is located outside nature reserves. The largest patch (c. 260 km^2) is on Mt Baigu, and other patches are located to the north-east of Taroko National Park, to the east of Yuli Wildlife Refuge, to the east and west of Guanshan Wildlife Refuge and to the west of Yushan National Park (Figs 1 & 3). We recommend evaluating the establishment of a nature reserve at Mt Baigu and expansion of other wildlife refuges and national parks. The sambar deer is a flagship species for conservation in Taiwan and the protection of its habitats would benefit other large mammals such as Reeves' muntjac *Muntiacus reevesi*, Formosan serow *Capricornis swinhoei* and black bear *Ursus thibetanus*. Our finding that distance to road and mean elevation are the most important factors determining habitat suitability is similar to the findings of Kushwaha et al. (2004), who suggested that sambar deer avoid direct contact with humans, preferring areas of higher elevation. Our map derived from ensemble forecasting shows that three highways separate potential habitat into five main patches. These highways were constructed c. 40–50 years ago. Traffic, human settlements, lights, noise, dogs and the presence of tourists along the roads cause disturbance to wild animals (Debeljak et al., 2001; Klar et al., 2008), and areas near the roads are vulnerable to poaching activity. Such human disturbance interrupts connectivity between patches separated by roads, and we hypothesize that gene flow between these patches has been limited in recent decades. The division of a species into small populations results in genetic characteristics being strongly influenced by inbreeding and genetic drift (Frankham, 1996). It may therefore be important to establish connections, such as bridges or tunnels at main crossing points along roads, between patches (Kuemmerle et al., 2010; Monterrubio-Rico et al., 2010). We recommend that a number of suitable habitat sites that are in close proximity to the three highways should be selected to monitor the population expansion.

Elevation is another important determinant of sambar deer habitat suitability; the species prefers areas of medium to high elevation. A study by Podchong et al. (2009) also indicated that geographical parameters affect sambar deer distribution. We suggest that the preference for higher elevations may be attributable in part to land exploitation at lower elevations. The sambar deer formerly occurred down to 300 m in Taiwan (Kano, 1940), and the bones of sambar deer have been found at low-elevation archaeological sites (Chen, 2000). Most areas of low elevation are now exploited and, because of dense human populations, there is no intact habitat available for the species in these areas. Forest cover and annual precipitation have been found to be important determinants of habitat suitability for the sambar deer (Kushwaha et al., 2004; Gormley et al., 2011) but these two variables did not influence the delimitation of habitat suitability in our modelling, possibly because their occurrence is correlated with elevation. Slope aspect is a predictor of the abundance of sambar deer (Forsyth et al., 2009) but we did not use this variable in our modelling because it is not suitable for use at a resolution of 1 km^2 grid cells. In our modelling of the habitat potentially suitable for the sambar deer, ensemble forecasting performed better than the five individual models. Although the value of AUC for the maximum entropy model was only 0.013 lower than that for ensemble forecasting, the prediction of ensemble forecasting was more comprehensive. Ensemble forecasting can produce more robust predictions of species characteristics (Araújo & New, 2007; Marmion et al., 2009; Thuiller et al., 2009; Oppel et al., 2012).

The location data used in this study were assimilated from many independent field surveys because we wanted to analyse potential deer habitats across the maximum possible range of areas. Use of the sampling methods of Forsyth et al. (2009) and Gormley et al. (2011) would provide data for more robust prediction models. However, the complex topography and low road density in the mountainous areas in Taiwan limit the application of such sampling methods, which would need to be modified before they could be used in any future studies. Our map of the potential distribution of the sambar deer in Taiwan and our indication of areas that are priorities for increased monitoring and/or protection provide a baseline for further research, and our modelling approach could be used elsewhere in the species' range. As a result of our studies and recommendations the national parks' administration is beginning a project to extensively monitor sambar deer population dynamics and habitat use.

Acknowledgements

We thank P.F. Lee for providing constructive comments at an early stage of the paper and for support with the environmental variables database. We thank C.Y. Lin for providing unpublished data, W. McShea for helpful discussions on the manuscript, and the reviewers for their useful comments. We are also grateful to the Taroko National Park and the Forestry Bureau of Taiwan for their funding of the field data collection.

References

ANDERSON, R.P., LEW, D. & PETERSON, A.T. (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162, 211–232.

ARAÚJO, M.B. & NEW, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, 22, 42–47.

BASKIN, L. & DANELL, K. (2003) *Ecology of Ungulates—A Handbook of Species in Eastern Europe and Northern and Central Asia*. Springer-Verlag, Berlin/Heidelberg, Germany.

BOITANI, L., SINIBALDI, I., CORSI, F., DE BIASE, A., CARRANZA, I.d.I., RAVAGLI, M. et al. (2008) Distribution of medium- to large-sized African mammals based on habitat suitability models. *Biodiversity and Conservation*, 17, 605–621.

BROTONS, L., THUILLER, W., ARAÚJO, M.B. & HIRZEL, A.H. (2004) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27, 437–448.

CEBALLOS, G. & EHRLICH, P.R. (2002) Mammal population losses and the extinction crisis. *Science*, 296, 904–907.

CENTRAL WEATHER BUREAU OF TAIWAN (2013) Http://www.cwb.gov.tw [accessed 6 July 2013].

Chen, K.T. (2000) On Taiwan mammalian faunas in different periods of time and related problems: the background materials for Taiwan zooarchaeological studies: I. *Bulletin of the Institute of History and Philology*, 71, 129–198.

Côté, S.D., Rooney, T.P., Tremblay, J.P., Dussault, C. & Waller, D.M. (2004) Ecological impacts of deer overabundance. *Annual Review of Ecology, Evolution, and Systematics*, 35, 113–147.

Debeljak, M., Dzeroski, S., Jerina, K., Kobler, A. & Adamic, M. (2001) Habitat suitability modelling for red deer (*Cervus elaphus* L.) in south-central Slovenia with classification trees. *Ecological Modelling*, 138, 321–330.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A. et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129–151.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49.

Forest Bureau of Taiwan (2009) Http://www.forest.gov.tw [accessed 6 July 2013].

Forsyth, D.M., McLeod, S.R., Scroggie, M.P. & White, M. (2009) Modelling the abundance of wildlife using field surveys and GIS: non-native sambar deer (*Cervus unicolor*) in the Yarra Ranges, south-eastern Australia. *Wildlife Research*, 36, 231–241.

Frankham, R. (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10, 1500–1508.

Freeman, E.A. & Moisen, G.G. (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, 217, 48–58.

Gavashelishvili, A. & Javakhishvili, Z. (2010) Combining radio-telemetry and random observations to model the habitat of near threatened Caucasian grouse *Tetrao mlokosiewiczi*. *Oryx*, 44, 491–500.

Gormley, A.M., Forsyth, D.M., Griffioen, P., Lindeman, M., Ramsey, D.S.L., Scroggie, M.P. & Woodford, L. (2011) Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology*, 48, 25–34.

Groom, M.J. (2006) *Threats to Biodiversity*. Sunauer Associates, Sunderland, USA.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29, 773–785.

Hirzel, A.H., Hausser, J., Chessel, D. & Perrin, N. (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, 83, 2027–2036.

Hirzel, A.H., Hausser, J. & Perrin, N. (2007) *Biomapper 4.0*. Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. Http://www.unil.ch/biomapper [accessed 31 August 2011].

Hosmer, D.W. & Lemeshow, S. (1989) *Applied Logistic Regression*. Wiley, New York, USA.

Huberty, C.J. (1994) *Applied Discriminant Analysis*. Wiley, New York, USA.

Kano, T. (1940) *Zoogeographic Studies of the Tsugitaka Mountains of Formosa*. Shibusawa Institute for Ethnographic Research, Tokyo, Japan.

Klar, N., Fernandez, N., Kramerschadt, S., Herrmann, M., Trinzen, M., Buttner, I. & Niemitz, C. (2008) Habitat selection models for European wildcat conservation. *Biological Conservation*, 141, 308–319.

Kuemmerle, T., Perzanowski, K., Chaskovskyy, O., Ostapowicz, K., Halada, L., Bashta, A.T. et al. (2010) European bison habitat in the Carpathian Mountains. *Biological Conservation*, 143, 908–916.

Kushwaha, S.P.S., Khan, A., Habib, B., Quadri, A. & Singh, A. (2004) Evaluation of sambar and muntjak habitats using geostatistical modelling. *Current Science*, 86, 1390–1400.

Lee, L.L., Lin, C.Y. & Chi, W.J. (2007) *Monitoring of Large Mammals and Population of Formosan Sambar Deer along the 2nd Section of Southern Central Ridge Trail in Yushan National Park*. Yushan National Park Headquarters, Nantou, Taiwan.

Lee, P.F., Liao, C.Y., Lee, Y.C., Pan, Y.H., Fu, W.H. & Chen, H.W. (1997) *An Ecological and Environmental GIS Database for Taiwan*. Council of Agriculture, Taipei, Taiwan.

Lee, P.F., Lue, K.Y. & Wu, S.H. (2006) Predictive distribution of hynobiid salamanders in Taiwan. *Zoological Studies*, 45, 244–254.

Leslie, D.M. (2011) *Rusa unicolor* (Artiodactyla: Cervidae). *Mammalian Species*, 43, 1–30.

Lin, L.K. (1997) *A Survey on the Fauna of the Alisan and Lulinsan Hardwood Nature Reserves*. Council of Agriculture, Taipei, Taiwan.

Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38, 921–931.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distribution*, 15, 59–69.

McLachlan, G.J. (2004) *Discriminant Analysis and Statistical Pattern Recognition*. Wiley, New York, USA.

McPherson, J.M., Jetz, W. & Rogers, D.J. (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, 41, 811–823.

Monterrubio-Rico, T.C., Renton, K., Ortega-Rodríguez, J.M., Pérez-Arteaga, A. & Cancino-Murillo, R. (2010) The endangered yellow-headed parrot *Amazona oratrix* along the Pacific coast of Mexico. *Oryx*, 44, 602–609.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A.F., Miller, P.I. & Louzao, M. (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biological Conservation*, 156, 94–104.

Pearson, R.G., Dawson, T.P. & Liu, C. (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, 27, 285–298.

Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brontons, L., McClean, C. et al. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33, 1704–1711.

Pei, K.K.C. (2004) *Mammal Survey in Ta-Hsueh-Shan Area, Shei-Pa National Park*. Shei-Pa National Park Headquarters, Miaoli, Taiwan.

Pei, K.K.C., Chou, C.H., Chen, M.T., Guo, Y.L. & Liu, Y.F. (2002) *Study on Terrestrial Mammals in the Kenting National Park (III)*. Kenting National Park Headquarters, Pingtung, Taiwan.

Pei, K.K.C., Chen, C.C. & Chen, M.T. (2003) *Monitoring of Medium-to-Large Sized Mammal Populations in Taroko National Park*. Taroko National Park Headquarters, Hualien, Taiwan.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.

Podchong, S., Schmidt-Vogt, D. & Honda, K. (2009) An improved approach for identifying suitable habitat of Sambar Deer (*Cervus unicolor* Kerr) using ecological niche analysis and environmental categorization: case study at Phu-Khieo Wildlife Sanctuary, Thailand. *Ecological Modelling*, 220, 2103–2114.

Porwal, M.C., Roy, P.S. & Chellamuthu, V. (1996) Wildlife habitat analysis for 'sambar' (*Cervus unicolor*) in Kanha National Park using remote sensing. *International Journal of Remote Sensing*, 17, 2683–2697.

Steinmetz, R., Chutipong, W., Seuaturien, N., Chirngsaard, E. & Khaengkhetkarn, M. (2009) Population recovery patterns of Southeast Asian ungulates after poaching. *Biological Conservation*, 143, 42–51.

Stockwell, D.R.B. & Peters, D. (1999) The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13, 143–158.

Su, H.J. (1992) Vegetation of Taiwan: altitudinal vegetation zones and geographical climatic regions. In *The Biological Resources of Taiwan: A Status Report* (ed. C.I. Peng), pp. 39–53. Institute of Botany, Academia Sinica, Taipei, Taiwan.

Teixeira, J., Ferrand, N. & Arntzen, J.W. (2001) Biogeography of the golden-striped salamander *Chioglossa lusitanica*: a field survey and spatial modelling approach. *Ecography*, 24, 618–624.

Thuiller, W., Araújo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L. & Lavorel, S. (2004) Biodiversity conservation: uncertainty in predictions of extinction risk. *Nature*, 430, 1–2.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369–373.

Timmins, R.J., Steinmetz, R., Sagar Baral, H., Samba Kumar, N., Duckworth, J.W., Anwarul Islam, M. et al. (2008) *Rusa unicolor*. In *The IUCN Red List of Threatened Species v. 2011.1*. Http://www.iucnredlist.org [accessed 31 August 2011].

Wang, Y. (1994) *The fauna survey at Chatianshan Nature Reserve*. Council of Agriculture, Taipei, Taiwan.

Wang, Y., Wang, C.C., Kuo, C.Y., Wu, H.J., Chen, S.C. & Tsai, C.C. (2002) *Status of Large Mammals in Taiwan (IV)*. Council of Agriculture, Taipei, Taiwan.

Wilson, D.E. & Reeder, D.M. (eds) (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference*. The Johns Hopkins University Press, Baltimore, USA.

Wu, H.Y. & Shi, J.D. (2006) *Monitoring the Artiodactyla Mammals at the Eastern Area in Yushan National Park*. Yushan National Park Headquarters, Nantou, Taiwan.

Wu, H.Y. & Yao, C.L. (2007) *Monitoring Mammal Activities and Mammal–Human Interactions in Nan-an/Bao-ai at the Eastern Area of Yushan National Park*. Yushan National Park Headquarters, Nantou, Taiwan.

Appendix

Descriptions of mathematical models

Logistic regression is a tool for analysing the effects of one or several independent variables, either discrete or continuous, on a dichotomic (presence/absence) or polychotomic dependent variable (Hosmer & Lemeshow, 1989). Logistic regression takes the following form:

$$\pi(x) = e^{g(x)}/\left(1 + e^{g(x)}\right) \text{ or } \pi(x) = 1/\left(1 + e^{-g(x)}\right)$$

where $\pi\ (x)$ represents the probability of occurrence of the target species and $g\ (x)$ is obtained using a regression equation of the form

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

where β_0 is a constant and $\beta_1, \beta_2, \ldots, \beta_p$ are the coefficients of independent variables $x_1, x_2, \ldots, x_p$, respectively (Hosmer & Lemeshow, 1989). Logistic regression analyses were performed using *SAS v. 9.0* (SAS Institute, Cary, USA).

Discriminant analysis is used to classify a set of observations into predefined classes that are based on a set of variables (McLachlan, 2004). It constructs a set of linear functions of the environmental variables, known as discriminant functions, whereby

$$L = b_1 x_1 + b_2 x_2 + \cdots + b_n x_n + c$$

where $b_1, b_2, \ldots, b_n$ are discriminant coefficients, $x_1, x_2, \ldots, x_n$ are the environmental variables and c is a constant. These discriminant functions are used to predict the class of a new observation with an unknown class. For a k class problem, k discriminant functions are constructed. Given a new observation, all of the k discriminant functions are evaluated and the observation is assigned to class i if the ith discriminant function has the highest value.

Ecological-niche factor analysis compares the distributions of the environmental variables in the presence dataset with those in the whole study area (Hirzel et al., 2002). This technique summarizes environmental variables into a few uncorrelated factors that explain most of the information. The output includes eigenvalues and factor scores. The first factor is the marginality factor, which describes the difference between the mean habitat in the study area and the species mean. The remaining factors are the specialization factors, which describe how specialized the species is with reference to the available habitat range in the study area (Hirzel et al., 2002). We performed this analysis using *Biomapper 4.0* (Hirzel et al., 2007). After computing the factor scores we used the algorithm of the medians to draw a habitat suitability map for sambar deer (Hirzel et al., 2002).

The genetic algorithm for rule-set production creates ecological niche models for species (Stockwell & Peters, 1999). The models describe environmental conditions under which a species should be able to maintain populations. The algorithm searches iteratively for non-random correlations between presence and environmental variables by using four types of rules: atomic, logistic regression, bioclimatic envelope and negated bioclimatic envelope. Predicted presence is defined by these rules. We used *Desktop GARP 1.1.6* (University of Kansas Center for Research, Kansas, USA) and followed the normal procedure for implementation. The output is a binary map; hence, we applied a modification of the best-subsets procedure described by Anderson et al. (2003). We ran 200 models and selected the 20 that had the highest predicted accuracy. The final prediction was produced by summing the

20 selected models, which yielded prediction values in the range 0–20.

Maximum entropy is a machine-learning technique that is based on the principle of maximum entropy (Pearson et al., 2004). It estimates the probability distribution of maximum entropy for each environmental variable across the study area with presence-only data (Pearson et al., 2004, 2006). This distribution is calculated with the constraint that the expected value of each environmental variable under this estimated distribution matches its empirical mean (Pearson et al., 2006). Habitat suitability maps were calculated by applying Maxent models to all grids in the study area, using a logistic link function to yield probability values ranging from 0 to 1. Maximum entropy performs well with small sample sizes (Elith et al., 2006; Hernandez et al., 2006). We developed our models using *Maxent v. 3.3.1* (Princeton University, Princeton, USA).

Biographical sketches

SHIH-CHING YEN studies the behaviour, ecology and management of sambar deer and sika deer. He is currently studying the habitat selection and space use of sambar deer, using radiotelemetry. YING WANG studies the ethology of birds and mammals and is involved in wildlife management and training aboriginal hunters as ecotourism guides. HENG-YOU OU studies species distribution modelling and spatial statistics.