International Scholars Journals

African Journal of Ecology and Ecosystems ISSN: 9428-167X Vol. 6 (2), pp. 001-005, February, 2019. Available online at www.internationalscholarsjournals.org © International Scholars Journals

Full Length Research Paper

Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia

Sunil Kumar[1]* and Thomas J. Stohlgren[2]

[1]1499 Campus Delivery, A204, NESB, Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499, USA.

[2]U.S. Geological Survey, Fort Collins Science Center, Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499, USA.

Accepted 21 November, 2018

Distribution data on threatened and endangered species are often sparse and clustered making it difficult to model their suitable habitat distribution using commonly used modeling approaches. We used a novel method called maximum entropy distribution modeling or Maxent for predicting potential suitable habitat for *Canacomyrica monticola* , a threatened and endangered tree species in New Caledonia, using small number of occurrence records (11). The Maxent model had 91% success rate (that is, a low omission rate) and was statistically significant. The approach presented here appears to be quite promising in predicting suitable habitat for threatened and endangered species with small sample records and can be an effective tool for biodiversity conservation planning, monitoring and management.

Key words: Biodiversity conservation, *Canacomyrica monticola*, hotspot, Maxent, New Caledonia, threatened and endangered species, small sample size.

INTRODUCTION

Prediction and mapping of potential suitable habitat for threatened and endangered species is critical for monitoring and restoration of their declining native populations in their natural habitat, artificial introductions, or selecting conservation sites, and conservation and management of their native habitat (Gaston, 1996). But distribution data on threatened and endangered species are often sparse (Ferrier et al., 2002; Engler et al., 2004) and clustered making commonly used habitat modeling approaches difficult.

Species distribution modeling tools are becoming increasingly popular in ecology and are being widely used in many ecological applications (Elith et al., 2006; Peterson et al., 2006). These models establish relationships between occurrences of species and biophysical and environmental conditions in the study area. A variety of species distribution modeling methods are available to predict potential suitable habitat for a species (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Elith et al., 2006; Guisan et al., 2007a,b; Wisz et al., 2008). However, comparatively few predictive models have been used for rare and endangered plant species (Engler et al., 2004) and there are fewer examples of studies using small sample sizes (For example, Pearson et al., 2007; Thorn et al., 2009). Most species distribution modeling methods are sensitive to sample size (Wisz et al., 2008) and may not accurately predict habitat distribution patterns for threatened and endangered species.

Our objectives were to: (1) predict suitable habitat distribution for threatened and endangered tree *Canacomyrica monticola* using a small number of occurrence records to inform conservation planning in New Caledonia; and (2) identify the environmental factors associated with *C. monticola*'s habitat distribution. We used species occurrence records, GIS (geographical information system) environmental layers (bioclimatic and topo-

*Corresponding author. E-mail: sunil@nrel.colostate.edu, sunilkrv@gmail.com. Tel.: 001-970-491-2302. Fax: 001-970-491-1965.

graphic), and the maximum entropy distribution modeling approach (Phillips et al., 2006) to predict potential suitable habitat for *C. monticola*.

MATERIALS AND METHODS Target

species and occurrence data

We obtained eleven occurrence records of *C. monticola* tree species from Herbert (2006); these records represent the total known distribution of the species. *C. monticola* is a threatened and endangered tree species, endemic to the pacific island group of New Caledonia. It grows in the patches of primary forest on ultramafic soils in the southern parts of Grande Terre, the main island of New Caledonia, one of the world's 25 biodiversity 'hotspots' defined by Myers et al. (2000). New Caledonia has a high level of species richness and endemism (Lowry1998; Jaffre et al., 2004; Lowry et al., 2004; Murienne, 2009) and 14.4% of its plant species are Red Listed by IUCN (IUCN, 2008; Munginger et al., 2008). *Canacomyrica* is phylogenetically and biogeographically interesting genus (Pete Lowry II; personal communication). Very little is known about *C. monticola*'s ecology and habitat distribution, and its habitat is under severe threat due to many factors including deforestation (mainly due to open-cast mining for nickel ore), invasive exotic species, fire, agriculture and grazing (Herbert, 2006; Pascal et al., 2008).

Environmental variables

We considered twenty five environmental variables as potential predictors of the *C. monticola* habitat distribution (Table 1) . These variables were chosen based on their biological relevance to plant species distributions and other habitat modeling studies (For example, Kumar et al., 2006; Guisan et al., 2007a, b; Pearson et al., 2007; Murienne et al., 2009). Nineteen bioclimatic variables (Nix, 1986), biologically more meaningful to define eco-physiological tolerances of a species (Graham and Hijmans 2006; Murienne et al., 2009), were obtained from WorldClim dataset (Hijmans et al., 2005; http://www.worldclim.org/bioclim.htm). Elevation (Digital Elevation Model; DEM) data were also obtained from the WorldClim website; 1 km spatial resolution. The DEM data were used to generate slope and aspect (both in degrees) using Environmental Systems Research Institute's ARC GIS version 9.2, 'Suface Analysis' function (ESRI, Redlands, California, USA). Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation continuous field (VCF) data representing percent tree cover, percent herbaceous covers, and percent bare cover (Hansen et al., 2003) were acquired from the Global Land Cover Facility (GLCF), University of Maryland (http://glcf.umiacs.umd.edu/data/vcf/) website. Soil data layer for New Caledonia (Murienne et al., 2008) was not used in the analyses because of its coarse resolution. All environmental variables were resampled to 1 km spatial resolution. All the variables were tested for multicollinearity by examining cross-correlations (Pearson correlation coefficient, r) among them based on 211 localities- 11 species occurrence records plus 200 randomly generated samples from the area. Only one variable from a set of highly cross-correlated variables ($r > 0.75$) was included in the model based on the potential biological relevance to the distribution of the species and the ease of interpretation. For example, MODIS tree cover and herbaceous cover were correlated ($r = - 0.87$, $P < 0.0001$), we dropped herbaceous cover and included tree cover. Thus, the number of predictor variables was reduced to ten (Table 1).

Modeling procedure

We used a novel modeling method called maximum entropy distribution or Maxent which has been found to perform best among many different modeling methods (Elith et al., 2006; Ortega-Huerta and Peterson, 2008), and may remain effective despite small sample sizes (Hernandez et al., 2006; Pearson et al., 2007; Papes and Gaubert, 2007; Wisz et al., 2008; Benito et al., 2009). Maxent is a maximum entropy based machine learning program that estimates the probability distribution for a species' occurrence based on environmental constraints (Phillips et al., 2006). It requires only species presence data (not absence) and environmental variable (continuous or categorical) layers for the study area. We used the freely available Maxent software, version 3.1 (http://www.cs.princeton.edu/~schapire/maxent/), which generates an estimate of probability of presence of the species that varies from 0 to 1, where 0 being the lowest and 1 the highest probability. The 11 occurrence records and 10 environmental predictors were used in Maxent to model potential habitat distribution for *C. monticola*. Since our sample size was low (< 20) we used only linear and quadratic features and maintained other settings as default (Phillips et al., 2004).

Testing or validation is required to assess the predictive performance of the model. Ideally an independent data set should be used for testing the model performance, however, in many cases this will not be available, a situation particular prevalent for threatened and endangered species. Therefore, the most commonly used approach is to partition the data randomly into 'training' and 'test' sets, thus creating quasi-independent data for model testing (Fielding and Bell, 1997; Guisan et al., 2003). However, this approach may not work with a small number of samples because the 'training' and 'test' datasets will be very small (Pearson et al., 2007). Therefore, we explicitly followed Pearson et al. (2007) and used a jackknife (also called 'leave-one-out') procedure, in which model performance is assessed based on its ability to predict the single locality that is excluded from the 'training' dataset (Pearson et al., 2007). Eleven different predictions were thus made with one of the occurrence records excluded in each prediction and the final potential habitat map was generated using all records (Figure 1). We used the *P* value program provided by Pearson et al. (2007) to test the significance of the model. The jackknife validation test required the use of a threshold to define 'suitable' and 'unsuitable' areas. We used two different thresholds, the 'lowest presence threshold' (LPT, equal to the lowest probability at the species presence locations), and a fixed threshold of 0.10; for more details see Pearson et al. (2007).

RESULTS AND DISCUSSION

The Maxent model predicted potential suitable habitat for *C. monticola* with high success rates (that is, low omission rates), 82% at LPT and 91% at threshold of 0.10, and was statistically significant in both cases ($P < 0.0001$ for LPT, and $P = 0.0025$ for threshold of 0.10). Most suitable habitat for *C. monticola* was predicted in the southern parts of the main island (south province) in New Caledonia (Figure 1), and its distribution is quite fragmented. The Maxent model's internal jackknife test of variable importance showed that 'temperature seasonality (standard deviation)', and 'precipitation seasonality (coefficient of variation)' were the two most important predictors of *C. monticola* 's habitat distribution (Figure 2; Table 1). These variables presented the higher gain (that is, contained most information) compared to other variables (Figure 2; Table 1). Using four arbitrarily defined probability classes, the high suitability class had an area of 752 km^2; medium-1,237 km^2; low- 4,178 km^2; and very

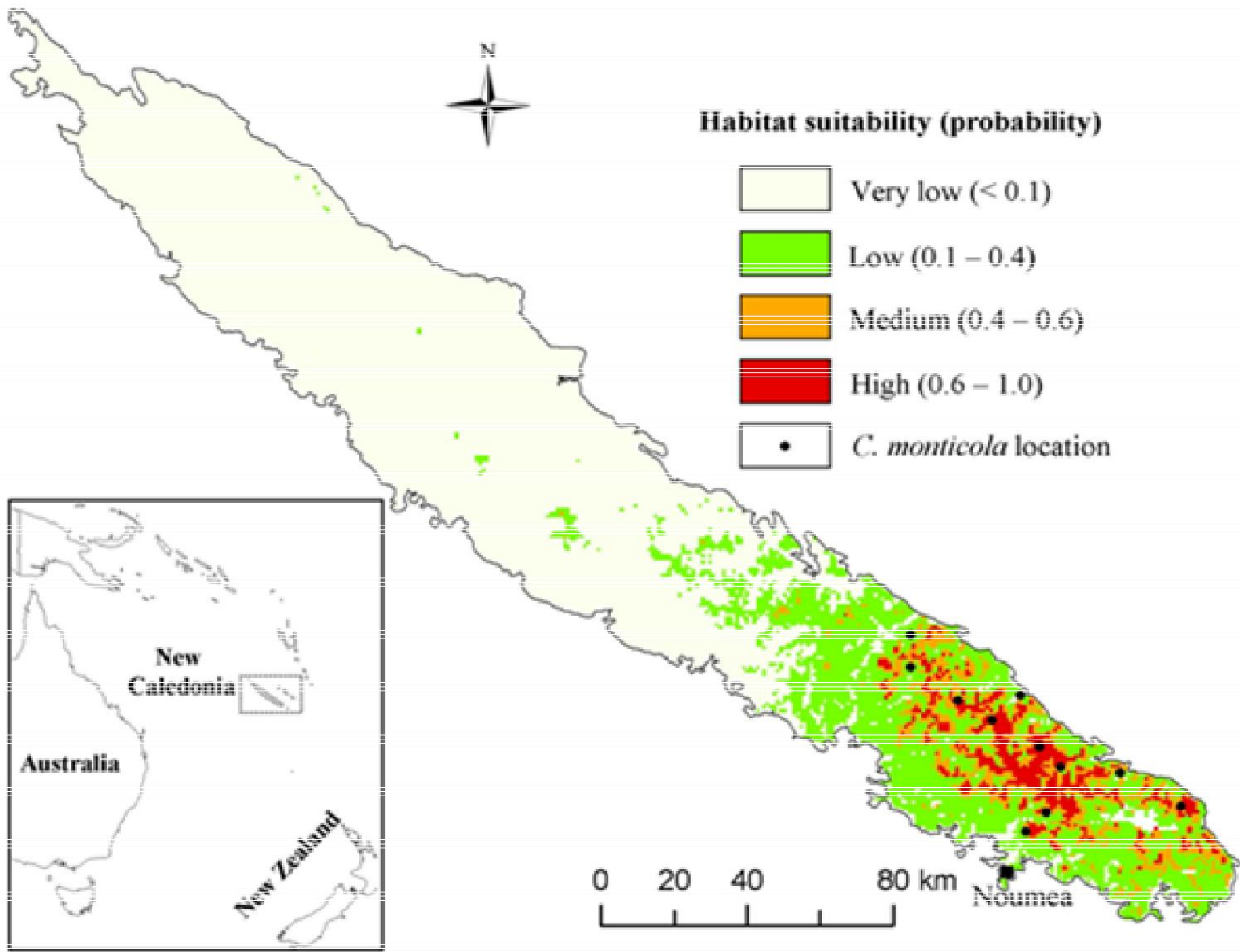

Figure 1. Predicted potential suitable habitat for *C. monticola* tree species on Grande Terre, the main island of New Caledonia.

Table 1. Selected environmental variables and their percent contribution in Maxent model for *Canacomyrica monticola* tree species in New Caledonia.

Environmental variable	**Percent contribution**	**Source/Reference**
Precipitation seasonality (coefficient of variation, Bio15)	51.2	WorldClim; Hijmans et al. 2005
Temperature seasonality (standard deviation x 100, Bio4)	25.8	WorldClim; Hijmans et al. 2005
Precipitation of coldest quarter (Bio19, degree C)	13.8	WorldClim; Hijmans et al. 2005
Aspect (degrees)	5.9	Generated in GIS
Mean temperature of wettest quarter (Bio8, degree C)	2.7	WorldClim; Hijmans et al. 2005
Precipitation of warmest quarter (Bio18, degree C)	0.5	WorldClim; Hijmans et al. 2005
MODIS Tree cover (%)	0.0	GLCF; Hansen et al. 2003
Elevation (m)	0.0	WorldClim; Hijmans et al. 2005
Slope (degrees)	0.0	Generated in GIS
MODIS bare cover (%)	0.0	GLCF; Hansen et al. 2003

Note: MODIS is Moderate Resolution Imaging Spectroradiometer; Bio 1, Bio2…Bio19 refer to 'Bioclimatic' variables obtained from WorldClim dataset- http://www.worldclim.org/GLCF is Global Land Cover Facility, University of Maryland, USA- http://glcf.umiacs.umd.edu/data/vcf/.

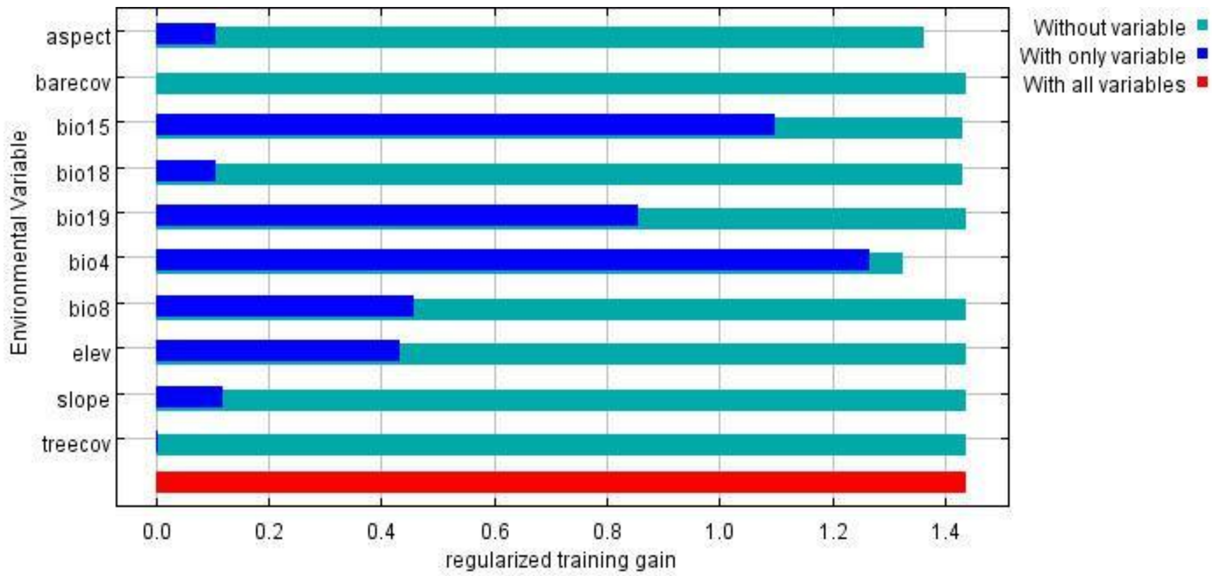

Figure 2. Results of jackknife evaluations of relative importance of predictor variables for *C. monticola* Maxent model. Note: 'elev is elevation; 'barecov' and 'treecov' are MODIS percent bare cover and tree cover; 'bio 4' is temperature seasonality, 'bio8' is mean temperature of wettest quarter, 'bio15' is precipitation seasonality, 'bio18' is precipitation of warmest quarter, and 'bio19' is precipitation of coldest quarter.

low-13,614 km^2 (Figure 1). The distribution of highly and moderately suitable areas appears to follow the distribution of ultramafic substrates in New Caledonia (Figure 1a in Grandcolas et al., 2008). The parts of the study area predicted in the 'very low' suitability class (probability < 0.10) can be interpreted as unsuitable for *C. monticola* (Figure 1). We also calculated total extent of occurrence (EOO, as defined by IUCN, 2001) of *C. monticola* based on the commonly used threshold of 0.5 (That is, the threshold above which the species is more likely to be present; Jimenez-Valverde and Lobo, 2007); it was estimated to be 1,305 km^2. This area is close to the 'manually measured' EOO of 1,420 km^2 for this species by Herbert (2006).

In this study we showed that the habitat distribution patterns for threatened and endangered plant species such as *C. monticola* can be modeled using a small num-ber of occurrence records and environmental variables using Maxent. This study provides the first predicted potential habitat distribution map for a plant species (*C. monticola*) in New Caledonia. Since Maxent is mapping the fundamental niche (different from occupied niche) of the species using bioclimatic variables the suitable habi-tat for *C. monticola* may be overpredicted in some areas (Pearson 2007; Murienne et al., 2009). However, the in-formation produced during this study is timely and highly relevant given the potential threats to *C. monticola* 's habi-tat and to overall biodiversity in New Caledonia due to nickel mining, anthropogenic burning, logging, and harm-ful invasive species (Herbert, 2006; Grandcolas et al., 2008). The potential habitat distribution map for *C. monti-cola* can help in planning land use management around its existing populations, discover new populations, identify top-priority survey sites, or set priorities to restore its natural habitat for more effective conservation. More research is needed to determine whether the existing protected areas adequately cover suitable habitat for *C. monticola*. The methodology presented here could be used for quantifying habitat distribution patterns for other threatened and endangered plant and animal species in other areas and may aid field surveys and allocation of conservation and restoration efforts.

ACKNOWLEDGEMENTS

We thank Natural Resource Ecology Laboratory at Colorado State University, Fort Collins, for logistical support. Funding was provided by U.S. Geological Survey and NASA grants (NRA-03-OES-03). Comments by Eric Guilbert, Pete Lowry II, Jerome Murienne, Bill Peterman, and two anonymous reviewers substantially improved the manuscript; to all we are grateful. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

REFERENCES

Benito BM, Martınez-Ortega MM, Munoz LM, Lorite J, Penas J (2009). Assessing extinction-risk of endangered plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses. Biodiv. Conserv. doi: 10.1007/s10531-009-9604-8.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006). Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

Engler R, Guisan A, Rechsteiner L (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. J. Appl. Ecol. 41: 263-274.

Ferrier S, Watson G, Pearce J, Drielsma M (2002) Extended statistical approaches to modeling spatial pattern in biodiversity: the north-east New South Wales experience. I. Species-level modeling. Biodiv. Conserv. 11: 2275–2307.

Fielding AH, Bell JF (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24: 38-49.

Gaston KJ (1996). Species richness: measure and measurement. Biodiversity: a biology of numbers and difference (ed. K.J. Gaston), Blackwell Science, Oxford pp. 77–113.

Graham CH, Hijmans RJ (2006). A comparison of methods for mapping species ranges and species richness. Glob. Ecol. Biogeogr. 15: 578–587.

Grandcolas P, Murienne J, Robillard T, Desutter-Grandcolas L, Jourdan H, Guilbert E, Deharveng L (2008). New Caledonia: A very old Darwinian island? Philo. Trans. Royal So. B-Biol. Sci. 363: 3309-3317.

Guisan A, Thuiller W (2005). Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8: 993–1009.

Guisan A, Zimmermann NE (2000). Predictive habitat distribution models in ecology. Ecol. Modell. 135: 147–186.

Guisan A, Graham CH, Elith J, Huettmann F and the NCEAS Species Distribution Modelling Group (2007a). Sensitivity of predictive species distribution models to change in grain size. Divers. Distrib. 13: 332-340.

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007b). What matters for predicting the occurrences of trees: techniques, data or species' characteristics? Ecol. Monogr. 77: 615-630.

Hansen M, DeFries RS, Townshend JRG, Carroll M, Dimiceli C, Sohlberg RA (2003). "Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm", Earth Interactions. 7(10): 1-15.

Hernandez PA, Graham CH, Master LL, Albert DL (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29: 773-785.

Herbert J (2006). Distribution, habitat and Red List status of the New Caledonian endemic tree *Canacomyrica monticola* (Myricaceae). Biodiv. Conserv. 15: 1459-1466.

Hijmans R J, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25: 1965-1978.

IUCN (2001). IUCN Red List Categories and Criteria: Version 3.1. IUCN Species Survival Commission, IUCN, Gland and Cambridge.

IUCN (2008). 2008 IUCN Red List of threatened species. URL http://www.iucnredlist.org

Jaffré T, Morat P, Rigault F, Veillon JM, Dagostini G (2004). Composition et caractéristiques de la flore indigène de la Nouvelle-Calédonie, Ed. 2. Nouméa: IRD.

Jimenez-Valverde A, Lobo JM (2007). Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecologica-Int. J. Ecol. 31: 361-369.

Kumar S, Stohlgren TJ, Chong GW (2006). Spatial heterogeneity influences native and nonnative plant species richness. Ecol. 87: 3186-3199.

Lowry PP II (1998). Diversity, endemism, and extinction in the flora of New Caledonia: A review. Pages 181-206 in Rare, Threatened, and Endangered Floras of Asia and the Pacific Rim, Institute of Botany, Academia Sinica Monograph Series No. 16, Taipei.

Lowry PP II, Munzinger J, Bouchet P, Géraux H, Bauer A, Langrand O, Mittermeier RA (2004). New Caledonia. Pages 193-197 in: Mittermeier RA, Gil P. R., Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux JL and da Fonseca GAB (eds.), Hotspots revisited: Earth's biologically richest and most threatened terrestrial ecoregions. CEMEX, Mexico City, Mexico.

Munzinger J, McPherson G, Lowry PP II (2008). A second species in the endemic New Caledonian genus *Gastrolepis* (Stemonuraceae) and its implications for the conservation status of high-altitude maquis vegetation: coherent application of the IUCN Red List criteria is urgently needed in New Caledonia. Bot. J. Linn. Soc. 157:775-783.

Murienne J, Pellens R, Budinoff RB, Wheeler WC, Grandcolas P (2008). Phylogenetic analysis of the endemic New Caledonian cock-roach *Lauraesilpha*. Testing competing hypotheses of diversification. Cladistics 24: 802-812.

Murienne J (2009). New Caledonia: biology. In Gillespie R. and Clagues D. (eds.). Pages 643-645; Encyclopedia of Islands; University of California Press, Berkeley.

Murienne J, Guilbert E, Grandcolas P (2009). Species' diversity in the New Caledonian endemic genera *Cephalidiosus* and *Nobarnus* (Insecta: Heteroptera: Tingidae), an approach using phylogeny and species' distribution modelling. Bot. J. Linn. Soc. 97:177-184.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000). Biodiversity hotspots for conservation priorities. Nature 403: 853–858.

Nix HA (1986). A biogeographic analysis of Australian Elapid snakes. In: Longmore R (Editor), Australian Flora and Fauna Series 8: 4-15.

Ortega-Huerta MA, Peterson AT (2008). Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. Revista Mexicana De Biodiversidad 79: 205-216.

Papes M, Gaubert P (2007). Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. Divers. Distrib. 13: 890-902.

Pascal M, De Forges BR, Le Guyader H, Simberloff D (2008). Mining and other threats to the New Caledonia biodiversity hotspot. Conserv. Biol. 22: 498-499.

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeo. 34: 102-117.

Pearson RG (2007). Species' distribution modeling for conservation educators and practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.

Peterson AT (2006). Uses and requirements of ecological niche models and related distributional models. Biodiv. Inform**.** 3: 59-72.

Phillips SJ, Dudik M, Schapire RE (2004). A maximum entropy approach to species distribution modeling. Proceedings of the 21st International Conference on Machine Learning. ACM Press, New York pp. 655–662.

Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190: 231-259.

Thorn JS, Nijman V, Smith D, Nekaris KAI (2009). Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (Primates: *Nycticebus*). Divers. Distrib. 15:289-298.

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. Divers. Distrib. 14: 763-773.