EFFECTS OF SAMPLE SIZE ON THE PERFORMANCE OF SPECIES DISTRIBUTION MODELS COUPLED WITH 3S TECHNOLOGY－A CASE STUDY OF *CASTANOPSIS CARLESSI*

Yao Yu Chiang[1], Nan-Chang Lo[2], Wei-I Chang[3] and Kai-Yi Huang[4]

[1] Dept. of Forestry, Chung-Hsing University, Taiwan, R.O.C.
[2]Experimental Forest Management Office, Chung-Hsing University, Taiwan, R.O.C.
[3] HsinChu F. D. O., Forest Bureau, Council of Agriculture, Taiwan, R.O.C.
[4]Dept. of Forestry, Chung-Hsing University, Taiwan, R.O.C.
[4]kyhuang@dragon.nchu.edu.tw (corresponding author)
[1234]250, Kuo-Kuang Rd., Taichung 402, Taiwan, R. O. C.

Proposed presenter: Yao-Yu Chiang
Suggested topics: Natural Resources—Forestry
Presentation Preference: Oral

Abstract: Species distribution model (SDM) has become an essential tool for decision making on forest management and ecological conservation. Many studies have applied 3S technology, including remote sensing (RS), geographic information system (GIS), global positioning system (GPS), and statistical methods to predict the potential habitat of a target species. Then information derived from modeling outcome can provide a basis of resource allocation and management for a decision-maker. However, samples used for SDM is often insufficient because field survey for collecting samples needs large amounts of cost and labor. Previous studies also indicated that inadequate sample size adversely affected model performance frequently. Hence, we attempted to examine how species sample size affects model performance in this study. *Castanopsis carlessi* (long-leaf chinkapins, LLC) trees grow widely over mountains in central Taiwan, especially those in our study area encompassing the Huisun Experimental Forest Station, but there is a minimum limit of tolerance on elevation (above 1,700 m). Because LLC's seeds have long been identified as an important food source for animals, the species has an important meaning and value in ecology. Accordingly, we chose LLC as target species for the study. GIS technique was applied to overlay the tree samples collected by GPS on the layers of elevation, slope, aspect, terrain position, and vegetation indices derived from SPOT-5 images, and then species presence-sample-only approaches, including BIOCLIM, DOMAIN and maximum entropy (MAXENT), were used for modeling the tree's potential habitat. We set four sizes (20, 40, 60, and 80) of species presence samples and chose *kappa* coefficient of agreement for model validation. The results showed that model's accuracy increases with the rise of sample size until it reaches a certain number *kappa* value may either approximate an asymptote or

gradually decline. BIOCLIM, DOMAIN, MAXENT performed best at sample size 60 (0.82), 80 (0.83) and 40 (0.89), respectively. However, the *kappa* value of BIOCLIM declined to 0.76 at size 80, while MAXENT still kept almost the same accuracy. Therefore, MAXENT may be an appropriate choice for SDM when sample size is small, especially for modeling rare species. In terms of the area of suitable habitat they predicted, MAXENT (2% of total area) is best suited for planning conservation site for endangered species, while BIOCLIM (3%) and DOMAIN (5%) man be useful to discover new populations of target species. Indeed, the elevation of LLCs over the study ranges from 1,700 m to 2,000 m, and almost no LLCs grow below 1,700 m or above 2,000 m. Hence, elevation was the most important variable in terms of contribution of predictor variables. The results of this study based on modeling the potential habitat of LLCs are consistent with ecological knowledge from traditional surveys.

Keyword: Species distribution model (SDM), *Castanopsis carlessi*, BIOCLIM, DOMAIN, maximum entropy (MAXENT), Species presence-sample-only approach