PREDICTING SPATIAL PATTERN OF *Rhododendron formosanum* BY MAXIMUM ENTROPY AND SUPPORT VECTOR MACHINE

Wei-Kai Lai[1], Nan-Jang 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: Wei-Kai Lai

Suggested topics: Geographic Information Systems (GIS)

Presentation Preference: Oral

ABSTRACT: Species distribution model (SDM), with the development of multivariate statistics and geo-spatial information technology, becomes an essential tool for studying spatial ecology and conservation planning. *Rhododendron formosanum* (red-stripe rhododendron, RSR) is an endemic species of Taiwan, and therefore the species was chosen as a target species for the study. There are two types of data (or algorithms) used in SDMs, "species presence sample only" (SPSO) and "species presence absence sample" (SPAS). Previous studies indicated that SPAS was better than SPSO because SPAS contains more information from species absence data. To check the argument, we chose four algorithms MAXENT, BIOCLIM, SVM, and ML, the first two are SPSO, and the last two are SPAS. GIS technique was used to overlay RSR samples obtained by GPS with layers of elevation, slope, aspect, terrain position, and vegetation indices derived from SPOT-5 images. Four models were developed by aforementioned algorithms to predict the spatial pattern of RSRs in the Huisun study area and to compare the performance of two types of methods. The results revealed that the mean *kappa* value of SPSO (0.65) was just slightly higher than that (0.67) of SPAS and the difference was not significant ($t = -0.52$, $p > 0.05$). SPAS method was not superior to SPSO even with more species absence information, and this result contradicted with outcome of previous studies. Besides, predictive accuracy varied with different algorithms. MAXENT (0.72) was the best, followed by SVM (0.67), and BIOCLIM (0.62) and ML (0.62) were the worst. Furthermore, MAXENT and SVM models can greatly reduce the area of field survey to 5% (856.8 ha) of the entire study area, and thereby saving both cost and labor. More importantly, the models merely based on

topographic variables could not extend RSR distribution pattern from one watershed to another one because the topographic attributes of Tong-Feng and Kuan-Dau watersheds in Huisun are quite different from each other. Future studies will attempt to incorporate high spatial resolution images, high spectral resolution images into a predictive model so that its accuracy can be improved and it can be applied on a larger scale.

Keywords: Species distribution model (SDM), Maximum entropy (MAXENT), Support vector machine (SVM), BIOCLIM, Maximum likelihood (ML), *Rhododendron formosanum*m.