Examing the Impact of Sampling Scheme and Environmental Variables on Spatial Extrapolation Performance of Mikania micrantha

Hao-Yuan Hung[1], Chin-Jin Kuo[1], Bao-Hua Shao[2], Nan-Chang Lo[3], Kai-Yi Huang[1]*

[1] Dept. of Forestry, National Chung-Hsing University

[2] Pu-Li Workstation, Nan-Tou Division Office, Forest Bureau, Council of Agriculture

[3] Experimental Forest Management Office, National Chung-Hsing University

Spatial extrapolation plays a crucial role in uncovering new populations of species that have been not previously found, making it highly meaningful in ecology. In this study, an unmanned aerial vehicle (UAV) was used to obtain four-date and five-band orthoimages of Mikania micrantha (bitter vine, BV). The flight missions were conducted over plots A and B, five kilometers apart in Zhushan, Formosa Island on January 27, April 6, August 23, and November 13, 2021. In addition, we utilized Trimble R12 receiver with precision in centimeter grade to measure BV as true values for model validation. The study aimed to evaluate the impact of different sampling designs and the combination of pure spectral imagery with topographic wetness index (TWI) and normalized difference vegetation index (NDVI) on spatial extrapolation performance for BV. Three sets of sampling design were utilized: (1) conducting individual plot classification using training data from it's owned. (2) performing spatial extrapolation from one plot with training data to another without training data, whose validation data there represent previously unidentified new populations. (3) combining the data from both plots and performing extrapolation. We used machine learning, support vector machines (SVM) and random forests (RF); deep learning U-net and DeepLabV3 to recognize the spatial patterns of BV. The results indicated that for set 1, four models were able to recognize the distribution of BV, with F1 scores exceeding 0.85. Among them, the inclusion of NDVI contributed the most to U-net, while the other variables resulted in a marginal increase or decrease in accuracy for all models. In set 2, TWI showed no significant differences due to the small size of two plots and yielded slight improvement in the extrapolation performance of SVM and RF. However, integrating NDVI led to contrasting outcomes. Extrapolating from plot A to plot B resulted in an accuracy decrease with F1 score dropping to 0.18. Conversely, extrapolating from B to A improved U-net and DeepLabV3, with an average F1 score of 0.83. This discrepancy can be attributed to the positive NDVI values in B due to containing more green leaves, while A exhibited negative NDVI values due to the significant increase in visible bands reflectance caused by the abundant white flowers of BV, so that extrapolated from A to B cannot grasp its features and have a negative effect. In set 3, combining all training data and incorporating NDVI significantly improved the accuracy of U-net and DeepLabV3, with the highest F1 score reaching 0.94. This result indicates that combining training data from two plots can reduce spatial heterogeneity and improved extrapolation performance. Therefore, training data should be taken using a larger number of smaller separate sampling areas. Furthermore, there are no one-size-fits-all variables combination that can achieve the best performance. Instead, it is necessary to use different combinations of variables to achieve optimal results. Future work will focus on exploring the potential of other deep learning models such as convolutional neural network (CNN) and environmental variables to improve the extrapolation performances of the models.

Keywords: spatial extrapolation, unmanned aerial vehicle (UAV), phenology, topographic wetness index (TWI), normalized difference vegetation index (NDVI)