

Plotting receiver operating characteristic and precision-recall curves from presence and background data

Wenkai Li¹ and Qinghua Guo²

¹Sun Yat-Sen University

²Peking University

March 23, 2021

Abstract

1. The receiver operating characteristic (ROC) and precision-recall (PR) plots have been widely used to evaluate the performances of species distribution models. Plotting ROC/PR curves requires a traditional test set with both presence and absence data (namely PA approach), but species absence data are usually not available in reality. Plotting ROC/PR curves from presence-only data while treating background data as pseudo absence data (namely PO approach) may provide misleading results. 2. In this study we propose a new approach to calibrate the ROC/PR curves from presence and background data with user-provided information on a constant c , namely PB approach. An estimate of c can also be derived from the PB-based ROC/PR plots given that a model with good ability of discrimination is available. We used three virtual species and a real aerial photography to test the effectiveness of the proposed PB-based ROC/PR plots. Different models (or classifiers) were trained from presence and background data with various samples sizes. The ROC/PR curves plotted by PA approach were used to benchmark the curves plotted by PO and PB approaches. 3. Experimental results show that the curves and areas under curves by PB approach are more similar to that by PA approach as compared with PO approach. The PB-based ROC/PR plots also provide highly accurate estimations of c in our experiment. 4. We conclude that the proposed PB-based ROC/PR plots can provide valuable complements to existing model assessment methods, and they also provide an additional way to estimate the constant c (or species prevalence) from presence and background data.

Hosted file

auc_pb_Mar092021.pdf available at <https://authorea.com/users/403517/articles/514924-plotting-receiver-operating-characteristic-and-precision-recall-curves-from-presence-and-background-data>