

特征选择和混合效应的加入改善了随机森林模型 对单木枯死的预测

马颖皓¹

¹ 北京林业大学, 北京 100083

为了更好的预测林木的枯死, 对森林演变规律进行分析。基于山西省第七期(2004 年)和第八期(2009 年)全国森林资源连续清查数据。分别使用两种传统模型方法: 逻辑斯蒂模型(LR)和广义混合效应模型(GLMM)和基于随机森林算法的四种方法: 基础随机森林(RF)、混合效应随机森林(MERF)、使用 GA 算法进行特征选择的随机森林模型(GARF)和使用 GA 算法进行特征选择的混合效应随机森林模型(GAMERF), 对山西省天然林四个主要树种组(桦木、栎木、松木和杨树)的单木枯死率进行建模, 并使用 SHAP 分析对机器学习模型进行解释。结果表明①四种基于 RF 模型的方法在四个树种中的准确率都优于传统模型, 这说明 RF 模型可以改善单木枯死率的预测。②在模型种加入随机效应可以有效的提升模型的性能; ③使用 GA 筛选变量的随机森林模型, 表现较基础模型更好, 并可以更好的发掘气候方面的变量, 通过计算 SHAP 值对机器学习模型进行解释, 发现变量解释都符合生物学理论。