Preliminary machine learning model for citrus greening disease (Huanglongbing-HLB) prediction in Colombia

Abstract

Citrus greening disease (Huanglongbing-HLB) is considered the most destructive citrus disease worldwide. Of the three species of Candidatus liberibacter associated with HLB, two have been recently reported in Latin America. The first report of HLB in Colombia was in March 2016. In this paper, a dataset was extracted for six departments in the northern zone of Colombia, where has been previously reported, applying image georeferencing with QGIS Software. Preliminary Random Forest and K-Nearest Neighbors (KNN) machine learning models were used in order to test and validate the obtained results, for disease monitoring and HLB incidence prediction. The performance of both models was also compared, obtaining a 100% AUC value with Random Forest model.