Using hyperspectral imagery and a multi-stage machine learning algorithm to distinguish infection symptoms caused by two xylem-limited pathogens

Poblete T[1], Navas-Cortes JA[2], Camino C[3], Calderon R[4], Hornero A[2, 5], Gonzalez-Dugo V[2], Landa BB[2], Zarco-Tajada PJ[1, 2]

[1] School of Agriculture and Food (SAF-FVAS) and Faculty of Engineering and Information Technology (IE-FEIT), University of Melbourne, 3010 VIC Melbourne, Australia
[2] Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Avda. Menéndez Pidal s/n, 14004 Córdoba, Spain
[3] European Commission, Joint Research Centre (JRC), 21027 Ispra (VA), Italy
[4] Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell AgriTech, Cornell University, Geneva, NY, USA
[5] Department of Geography, Swansea University, SA2 8PP Swansea, United Kingdom

Two major threats affecting olives groves are the bacterium *Xylella fastidiosa* (Xf) and the soil borne fungus *Verticillium dahliae* (Vd) Kleb. Both pathogens block the water flow along the xylem vessels, triggering similar physiological symptoms which are also easily confounded with water-stress induced responses. Previous remote sensing studies have attempted to independently detect Xf and Vd infections using hyperspectral and thermal imagery, but the potential discrimination between both infections has not been addressed yet. In this study, a multistage machine learning (ML) methodology is proposed to detect and discriminate the spectral changes caused by Xf and Vd infection. The multi-stage ML approach assessed spectral plant traits retrieved by physical model inversion such as leaf pigment content (carotenoids, Cx+c, chlorophyll a+b, Ca+b and anthocyanins, Anth.), canopy structural parameters (leaf area index, LAI and the leaf inclination distribution function, LIDFa), solar-induced fluorescence (SIF@760), the thermal-based Crop Water Stress Index (CWSI) and narrow-band hyperspectral indices. On each stage, key indicators were used to i) detect infection by each of the two pathogens accounting for the importance of the spectral traits, ii) reduce the detection uncertainties due to the confounding physiological responses triggered by water stress, and iii) discriminate between Vd and Xf infections. Results showed that SIF@760, Anth. and the Normalized Phaeophytinization Index (NPQI) were the most relevant traits to identify trees showing symptoms confounded with water stress. Using a Vd+Xf cross-infected balanced dataset the multi-stage ML approach yielded overall accuracies that exceeded 92% (kappa=0.8). The most sensitive spectral plant-trait indicators to discriminate between both pathogen infections will be discussed, describing operational aspects for the large-scale monitoring of olive orchards in geographic regions potentially affected by both pathogens.

Keyword: Hyperspectral, thermal, machine learning, disease-differentiation