ISPRS Journal of Photogrammetry and Remote Sensing 179 (2021) 133–144

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Discriminating *Xylella fastidiosa* from *Verticillium dahliae* infections in olive trees using thermal- and hyperspectral-based plant traits

T. Poblete [a], J.A. Navas-Cortes [b], C. Camino [c], R. Calderon [d], A. Hornero [e,b], V. Gonzalez-Dugo [b], B.B. Landa [b], P.J. Zarco-Tejada [a,b,*]

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

ARTICLE INFO

Keywords:
Hyperspectral
Thermal
Machine learning
Plant traits
Verticillium dahliae
Xylella fastidiosa

ABSTRACT

Globally, olive (*Olea europaea* L.) productivity is threatened by plant pathogens, particularly the fungus *Verticillium dahliae* (*Vd*) and the bacterium *Xylella fastidiosa* (*Xf*). Infections by these pathogens restrict water and nutrient flow through xylem, producing a similar set of symptoms that can also be confounded with water stress. Conventional *in situ* monitoring techniques are time consuming and expensive, necessitating the development of large-scale detection methods. Airborne hyperspectral and thermal imagery have been successfully used to detect both *Xf* and *Vd* infection symptoms independently, i.e., when only one of the two diseases is present. Nevertheless, the discrimination of *Vd* from *Xf* infections in contexts where both pathogens are present has not been addressed to date. This study proposes a three-stage machine learning algorithm to distinguish *Vd* infections from *Xf* infections, using a series of datasets from 27 olive orchards affected by *Xf* and *Vd* outbreaks in Italy and Spain between 2011 and 2017. Plant traits were derived from airborne hyperspectral and thermal imagery, including physiological indices from radiative transfer model inversion, Solar-induced Fluorescence emission ($SIF_{@760}$), the Crop Water Stress Index (CWSI), and a selection of narrow–band hyperspectral indices. Several distinct spectral traits successfully discriminated *Xf* from *Vd* infections. The three-stage method generated a false-positive rate of 9%, an overall accuracy (OA) of 98%, and a kappa coefficient (κ) of 0.7 when identifying *Vd* infections using a mixed *Vd* + *Xf* dataset. When identifying *Xf* infections, the false-positive rate was 4%, the OA was 92%, and κ was 0.8. These results indicate that hyperspectral and thermal traits can be used to discriminate *Xf* from *Vd* infection caused by the two xylem–limited pathogens that trigger similar visual symptoms.

1. Introduction

Olive trees are cultivated on over 11.5 million ha of land worldwide, with most of the production occurring in the Mediterranean Basin (Rodríguez-Cohard et al., 2020). Regions such as the USA, Asia, Oceania (FAO, 2008), and, more recently, other areas in the southern hemisphere also produce olives (Torres et al., 2017). Over 100 pests and pathogens infect olive trees (*Olea europaea*), reducing yields and increasing total production costs (Fernández-Escobar et al., 2013). Two particularly destructive pathogens include the soil-borne fungus *Verticillium dahliae* (*Vd*) Kleb (Tsror, 2011) and the gram-negative bacterium *Xylella fastidiosa* (*Xf*) (EFSA, 2019).

Vd causes Verticillium wilt (VW) (Jiménez-Díaz et al., 2012), the main soil-borne fungal disease threatening olive production worldwide (Jiménez-Díaz et al., 2012). First described in Italy in 1946, VW is now present in many Mediterranean countries, including Spain, Italy, Greece, France, Malta, Turkey, Morocco, Jordan, Algeria, Israel, Iran, and Tunisia. VW has also been reported in Argentina, the USA, Australia, and Peru, with more than 80,000 ha affected globally (Jiménez-Díaz et al., 2012); López-Escudero and Mercado-Blanco, 2011). There are two main pathotypes of *Vd*: defoliating (D) and non-defoliating (ND). The D pathotype is more virulent than the ND pathotype and can be lethal for infected trees. The first report of the D pathotype in Spain was in an intensive cotton field, but it has since spread across olive orchards in Andalusia, where it is now the predominant *Vd* pathotype, causing significant yield losses and tree mortality (Navas-Cortés et al., 2008).

* Corresponding author.

https://doi.org/10.1016/j.isprsjprs.2021.07.014
Received 23 June 2021; Accepted 26 July 2021
Available online 4 August 2021

Xf is one of the most devastating plant pathogens of olive and is categorized as a quarantine plant pathogenic bacterium in the European Union (Schneider et al., 2020). *Xf* infections cause olive quick decline syndrome, characterized by leaf scorch and desiccation of twigs and branches. *Xf*-infected olive trees were first reported in California in 2003 (Wong et al., 2003) and then in Apulia, the main olive-growing area of Italy, in 2013 (Saponari et al., 2013). Since then, *Xf* infections have been identified in several European countries, infecting about 100 host plant species across Tuscany (Italy), Corsica and the PACA region (France), Alicante and the Balearic Islands (Spain), and Northern Portugal (EFSA, 2020). Countries in the Mediterranean basin in Europe (EFSA, 2019) and in the Middle East and Africa, including Lebanon, Turkey, Morocco, and Tunisia, are the most vulnerable to *Xf* outbreaks (Frem et al., 2020).

Both *Vd* and *Xf* are xylem-invading pathogens that colonize the host through vascular tissue and eventually block water flow through the xylem. *Vd* enters through the root system, reaches the xylem, and moves within the vascular stream. Symptoms appear when the fungus emerges, where it blocks vessels and colonizes neighboring tissues (Klosterman et al., 2009). *Xf* infections gradually block the xylem, leading to a reduction of sap flow due to bacterial growth and plant physiological responses (Sicard et al., 2018). Similar symptoms are triggered for both xylem-limited pathogens in the canopies of infected plants, including foliar discoloration, wilting of apical shoots, dieback of twigs and branches, and general decline (Carlucci et al., 2013). Symptoms caused by the two pathogens may also be confounded with generic water stress responses, due to the xylem-limiting nature of the infection processes (Hopkins, 1989; Klosterman et al., 2009).

Vd and *Xf* infections cannot be discriminated visually *in situ* due to their similar effects on host trees. Accurate identification requires laboratory testing using molecular diagnostics that are laborious and expensive (Gramaje et al., 2013). Consequently, there is a need for more cost-effective approaches to detect and discriminate infected plants at large spatial scales and within reasonable time frames. Such methods could enable eradication of incipient *Xf* outbreaks (Almeida, 2016) and reduce the serious damage caused by *Vd* worldwide. Remote sensing techniques have great potential for efficiently detecting early disease symptoms in olive trees (Martinelli et al., 2015). Calderón et al. (2013) successfully used hyperspectral and thermal sensors to detect *Vd* infections in olive trees, finding that the thermal Crop Water Stress Index (CWSI) (Idso et al. 1981), Solar-induced Chlorophyll Fluorescence (SIF) (Plascyk, 1975), and the blue-region spectral indices BGI1 (Zarco-Tejada et al., 2005), BRI1 (Zarco-Tejada et al., 2012), and B (Calderón et al., 2013) were the most sensitive for detecting *Vd*-induced symptoms at the early stages of infection.

Calderón et al. (2013) also established that several spectral indices were useful for assessing disease severity at advanced stages of disease development (i.e. for damage-mapping after symptom emergence). The indicators of advanced disease severity included the photochemical reflectance index (PRI) (Gamon et al., 1992) and variants of PRI, such as PRI_{515} (Hernández-Clemente et al., 2011), the Transformed Chlorophyll Absorption Reflectance Index (TCARI) (Haboudane et al., 2002), and the Healthy Index (HI) (Mahlein et al., 2013). A subsequent study by the same authors (Calderón et al., 2015) conducted over a large area comprising infected orchards differing in soil and crop management characteristics showed that the air-canopy temperature difference (Tc-Ta); SIF; PRI_{515}; HI; structural indices such as the Renormalized Difference Vegetation Index (RDVI), the Modified Triangular Vegetation Index ($MTVI_1$), the Modified Simple Ratio (MSR), the Optimized Soil-Adjusted Vegetation Index (OSAVI), and the Enhanced Vegetation Index (EVI); the chlorophyll indices TCARI, Gitelson & Merzlyak (GM_1), and Pigment Specific Simple Ratio ($PSSR_b$); and the carotenoid index (R_{515}/R_{570}) were the most sensitive spectral indicators for detecting *Vd* infections at both early and advanced stages of the infection. They reported an overall accuracy (OA) of 79.2% and a kappa coefficient (κ) of 0.49 when detecting *Vd*-infected olive trees across all severity levels. The authors of both studies (Calderón et al., 2013; Calderón et al., 2015) suggested that Tc-Ta and SIF were the most important indicators for detecting *Vd* infection symptoms at the early stages of disease development, both at the orchard scale and over larger areas. They reasoned that these thermal and fluorescence indicators of early infection were more robust than spectral reflectance indices to variability in the orchard management strategies typically encountered when monitoring large areas.

The first study to evaluate early detection of *Xf* symptoms using hyperspectral remote sensing imagery (Zarco-Tejada et al., 2018) found that a combination of spectral plant traits retrieved from radiative transfer model (RTM) inversion, narrow-band hyperspectral indices, and thermal traits successfully identified *Xf*-infected olive trees in Apulia, Italy. The authors demonstrated that the blue-region Normalized Phaeophytinization Index (NPQI) (Barnes et al., 1992; Peñuelas et al., 1995), the thermal CWSI, anthocyanin pigment content (Anth), and SIF were the most sensitive spectral-based plant traits for successfully discriminating asymptomatic from symptomatic *Xf*–infected trees. They achieved an OA of 80% and a κ of 0.61. Later, Poblete et al. (2020) compared the sensitivity of hyperspectral and multispectral bandsets for detecting *Xf* infections in an operational context, assessing the contribution of CWSI and SIF. They found that the thermal CWSI was crucial for detecting *Xf* and demonstrated that large-scale *Xf* detection is feasible with multispectral and thermal cameras, provided that more than six narrow bands (i.e., 10 nm bandwidths) centered at specific spectral regions are captured, including (in order of importance) 400, 423, 601, 667, 713, 534, 478, 769, 760, and 735 nm. The study demonstrated that the blue spectral region (i.e., bands within the 400–450 nm spectral range used to derive the NPQI index) and the thermal region (used to derive CWSI) were the most critical spectral regions to detect *Xf* symptoms in olive. Using six 10-nm bandwidth spectral bands coupled with CWSI, their OA and κ coefficients were 74% and 0.36, respectively. After CWSI and NPQI, the most sensitive indices were those related to carotenoid and xanthophyll pigment dynamics (DC_{axbc}); the carotenoid reflectance index (CRI_{700M}); PRI derivatives such as PRI_{M1}, PRI_{M4}, PRI_n, and PRI × CI; and chlorophyll indices such as Vogelmann (VOG_2) and TCARI normalized by the OSAVI index in the form TCARI/OSAVI.

Although both *Xf*- and *Vd*-induced symptoms have been successfully detected independently using hyperspectral and thermal remote sensing techniques, discriminating *Xf*- from *Vd*–induced symptoms when both pathogens are present has not been attempted to date. Both *Xf* and *Vd* are vascular pathogens that cause similar visual symptoms and may therefore be confounded when monitoring large outbreaks. *Xf*- and *Vd*–induced symptoms may also be confused with plant responses to abiotic stress, i.e., nutrient or water limitation. As both *Vd* and *Xf* continue to spread globally, detecting and discriminating these infections via remote sensing is increasingly important.

Machine learning techniques may be able to distinguish subtle physiological differences between pathogen-induced symptoms, particularly when combined with hyperspectral and thermal image data. In this study, we propose a multi-stage method to distinguish *Vd* and *Xf* infections in olive trees. We combined airborne hyperspectral datasets collected from *Xf*- (Zarco-Tejada et al., 2018; Poblete et al., 2020) and *Vd*-infected olive orchards (Calderón et al., 2013; 2015) to assess this method. The remainder of this manuscript is organized as follows: in Section 2, we describe the study areas in detail. We then describe airborne hyperspectral and thermal imagery collection, field assessments, and how *Vd* and *Xf* datasets were combined. At the end of Section 2, we outline the multi-stage machine learning methodology for identifying the diseases using the combined datasets. In Section 3, we present and discuss our results. Finally, we highlight the main conclusion in Section 4.

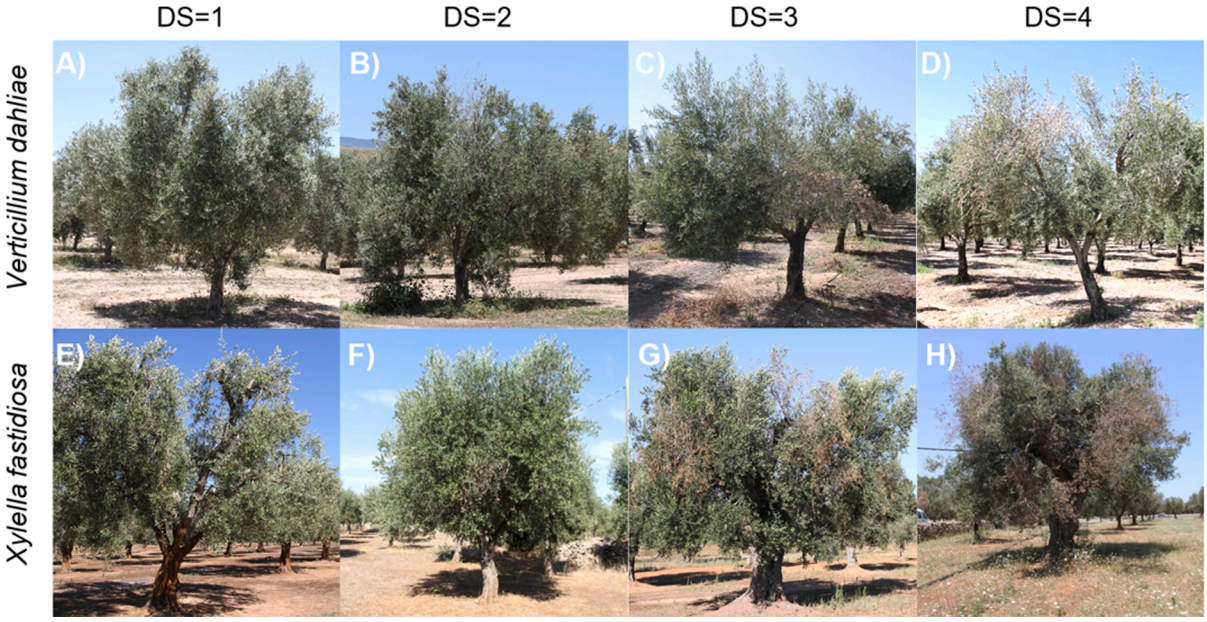

Fig. 1. Disease severity (DS) evaluation. Progressive symptom development in olive trees infected by *Verticillium dahliae* (A–D) and *Xylella fastidiosa* (E–H).

2. Material and methods

2.1. Study sites and Xf and Vd dataset description

Vd data were derived from 11 olive orchards in Southern Spain that were monitored in 2011 and 2013 (full description of the dataset can be found in Calderón et al. (2013) and Calderón et al. (2015)). The presence of *Vd* in these orchards was confirmed using a *Vd*-specific PCR assay (Mercado-Blanco et al. 2003). The D pathotype of *Vd* was prevalent in all the monitored olive orchards. Orchards were distributed across two study sites. The first study site was located in Castro del Rio (Cordoba, Spain) and contained commercial olive orchards of cv. Picual planted at a spacing of 6 X 4 m^2 (Calderón et al., 2013). Visual assessment of the disease incidence (DI) and severity (DS) was performed on a scale of 0–4 based on foliar symptoms and percentage of incidence (Fig. 1A–D). A total of 1,878 olive trees were assessed, among which 1,569 were asymptomatic and 283 were symptomatic (77% DS = 1, 16% DS = 2, 4% DS = 3, and 3% DS = 4). The second *Vd*-infected study site was located in Écija (Seville, Spain). Orchards comprised cv. Picual and cv. Hojiblanca, planted at densities of between 123 and 357 trees per ha (Calderón et al., 2015). Visual assessments of DI and DS were performed for 5,223 olive trees; 5,040 trees were asymptomatic and 183 were symptomatic, of which 61% had DS = 1, 22% had DS = 2, 12% had DS = 3, and 5% had DS = 4. The extension of the *Vd*-infected olive orchards used in this study ranged between 1.69 and 7.28 ha each.

Xf data were derived from 15 olive orchards in Puglia, southern Italy, monitored in 2016–2017 (a full description of the dataset can be found in Zarco-Tejada et al. (2018) and Poblete et al. (2020)). The presence of *Xf* was confirmed using an *Xf*-specific quantitative PCR (qPCR) assay (Harper et al., 2010). Orchards were composed of cv. Cellina di Nardò and cv. Ogliarola Salentina, ranging between 0.44 and 6.02 ha each. Visual assessments were performed on 7,296 olive trees; 4,045 were asymptomatic and 3,251 were symptomatic (Fig. 1E–H). Of the symptomatic trees, 45% had DS = 1, 41% had DS = 2, 11% had DS = 3, and 3% had DS = 4.

The *Vd* dataset was used in Calderón et al. (2013; 2015), and the *Xf* dataset was used in Zarco-Tejada et al. (2018) and Poblete et al. (2020) to assess the spectral plant traits sensitive to each disease. In this study, *Xf* and *Vd* datasets were combined, therefore generating mixed *Xf* + *Vd* datasets containing asymptomatic, *Vd*-infected, and *Xf*–infected trees. Using this mixed *Xf* + *Vd* dataset, specificity and sensitivity tests were conducted to assess the accuracy when discriminating *Xf* from *Vd* symptoms. The sensitivity test was performed by applying *Xf* and *Vd* discrimination models on datasets containing asymptomatic/symptomatic trees from the non-target disease (the *Xf* model tested on the *Vd* dataset and vice versa). The specificity test assessed the performance of each discrimination model using data derived from symptomatic and non-symptomatic trees from the non-target disease (95%) and symptomatic trees infected by the target infection (5%). More details regarding the proposed methodology and the construction of the datasets used to perform the sensitivity and specificity tests can be found in Section 2.3. To account for non-disease-induced water stress in *Xf* and *Vd* detection, hyperspectral and thermal datasets acquired in pathogen-free orchards were used as reference.

2.2. Hyperspectral and thermal airborne campaigns

Airborne campaigns were performed over every study site to acquire high-resolution thermal and hyperspectral imagery (Fig. 2A,B and C,D for *Vd*- and *Xf*-infected study sites, respectively). Flights were concurrent with field data collection. A FLIR Systems SC655c (USA) uncooled microbolometer-based detector thermal camera was used, sensitive to the spectral range 7.5–14 μm, with a resolution of 640 × 480 pixels and a focal length of 24.6 mm f/1.0. Thermal vicarious calibration was performed by measuring soil temperature as described in Calderón et al. (2013).

A Headwall Photonics (Fitchburg, MA, USA) VNIR linear hyperspectral imager was used to collect hyperspectral imagery. The imager is sensitive to the spectral range 400–885 nm and collects 260 bands with 6.4-nm full-width at half maximum (FWHM). The sensor has an 8-mm focal length that allows an angular field of view of 50°. Radiometric calibration was performed using a CSTM-USS-2000C integrating sphere (LabSphere, North Sutton, NH, USA). The Simple Model of Atmospheric Radiative Transfer of Sunshine (SMARTS) was used for atmospheric correction (Gueymard, 2001). A portable weather station (Transmitter PTU30, Vaisala, Helsinki, Finland) and MICROTOPS II sunphotometer (Solar LIGHT Co., Philadelphia, PA, USA) were used to measure meteorological parameters and aerosol optical properties at the time of hyperspectral and thermal acquisitions. The calibrated and atmospherically corrected hyperspectral data were ortho-rectified using IMU and

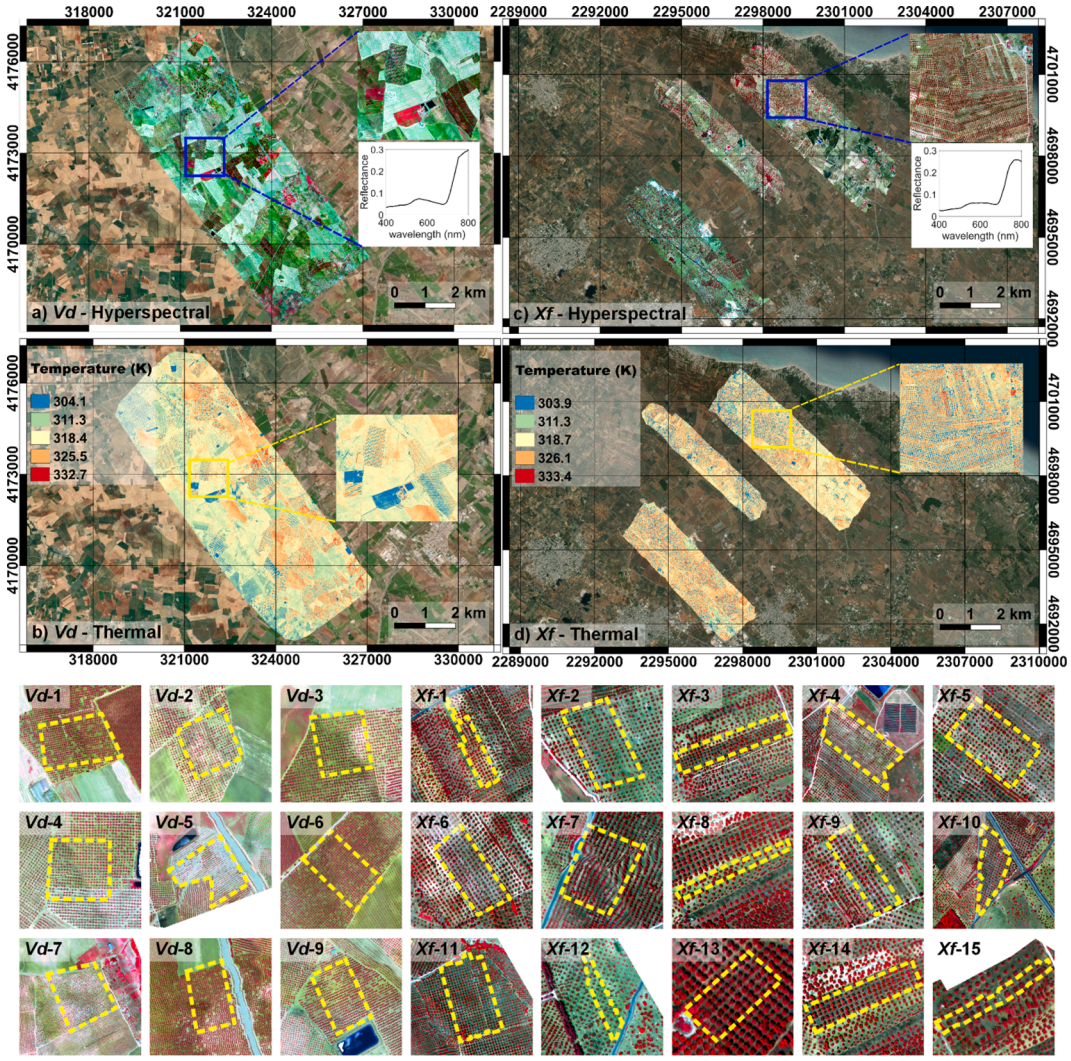

Fig. 2. Airborne hyperspectral and thermal imagery acquired over *Verticillium dahliae* (*Vd*)– A,B) and *Xylella fastidiosa* (*Xf*)–affected study sites (C,D). Hyperspectral (A) and thermal (B) imagery obtained over the *Vd* study site: *Vd*-1 (3.01 ha), *Vd*-2 (2.28 ha), *Vd*-3 (1.69 ha), *Vd*-4 (2.63 ha), *Vd*-5 (3.34 ha), *Vd*-6 (3.40 ha), *Vd*-7 (7.28 ha), *Vd*-8 (3.17 ha), and *Vd*-9 (2.54 ha). Hyperspectral (C) and thermal (D) imagery obtained over the *Xf* study sites: *Xf*-1 (1.42 ha), *Xf*-2 (2.39 ha), *Xf*-3 (1.46 ha), *Xf*-4 (6.02 ha), *Xf*-5 (2.53 ha), *Xf*-6 (1.67 ha), *Xf*-7 (1.36 ha), *Xf*-8 (0.54 ha), *Xf*-9 (1.16 ha), *Xf*-10 (3.41 ha), *Xf*-11 (2.92 ha), *Xf*-12 (0.47 ha), *Xf*-13 (0.52 ha), *Xf*-14 (1.52 ha), and *Xf*-15 (0.44 ha).

Table 1
Values and ranges used for model inversion and look-up table generation for the PRO4SAIL (PROSPECT-D + 4SAIL) radiative transfer model.

Parameter	Abbreviation	Value/range
Chlorophyll $a + b$ content [μg/cm^2]	C_{a+b}	10–70
Carotenoid content [μg/cm^2]	C_{x+c}	0–20
Anthocyanin content [μg/cm^2]	Anth	0–7.5
Dry matter content [g/cm^2]	C_m	0.012
Water content [g/cm^2]	C_w	0.009
Mesophyll struct. Coef.	N	1–2.5
Leaf area index [m^2/m^2]	LAI	0.3–5
Leaf Inclination Dist. Func. [deg.]	$LIDF_a$	0–90
Hot spot parameter	h	0.01
Soil reflectance	R_{soil}	PRO4SAIL dry soil spectra
Observer angle [deg.]	tt_o	0
Sun zenith angle [deg.]	tt_s	53.75[a], 66.9[b]
Relative azimuth angle [deg.]	ψ	0

[a] tts used for the Verticillium dahliae study sites.
[b] tts used for the Xylella fastidiosa study sites.

GPS data captured during the flights using the Parametric Geocoding & Ortho-rectification for Airborne Optical Scanner Data software (PARGE, ReSe Applications Schläpfer, Wil, Switzerland).

Pure tree-crown temperature and reflectance were obtained by performing Niblack's thresholding method (Niblack, 1986) and Sauvola's binarization techniques (Sauvola and Pietikäinen, 2000) to isolate tree crowns from the background and remove soil effects and within-crown shadows. The pure tree-crown temperature was used to calculate the CWSI (Idso et al., 1981) and to assess changes in the transpiration rates linked to *Xf*- (Zarco-Tejada et al., 2018) and *Vd*-induced (Bruno et al., 2020; Calderón et al., 2013) physiological changes. Pure tree-crown radiance and reflectance were used to i) quantify sun-induced chlorophyll fluorescence at 760 nm ($SIF_{@760}$) using the O_2-A *in–filling* Fraunhofer line depth (FLD) method (Plascyk, 1975); ii) calculate Narrow Band Hyperspectral Indices (NBHIs) (Table S1 of the Supplementary Material); and iii) retrieve leaf biochemical constituents and canopy structural parameters using the PRO4SAIL RTM that couples PROSPECT-D (Féret et al., 2017) and 4SAIL (Verhoef et al., 2007) models.

Leaf pigments such as carotenoids (C_{x+c}), chlorophyll $a + b$ (C_{a+b}), and anthocyanins (Anth) were retrieved from PROSPECT-D. Canopy

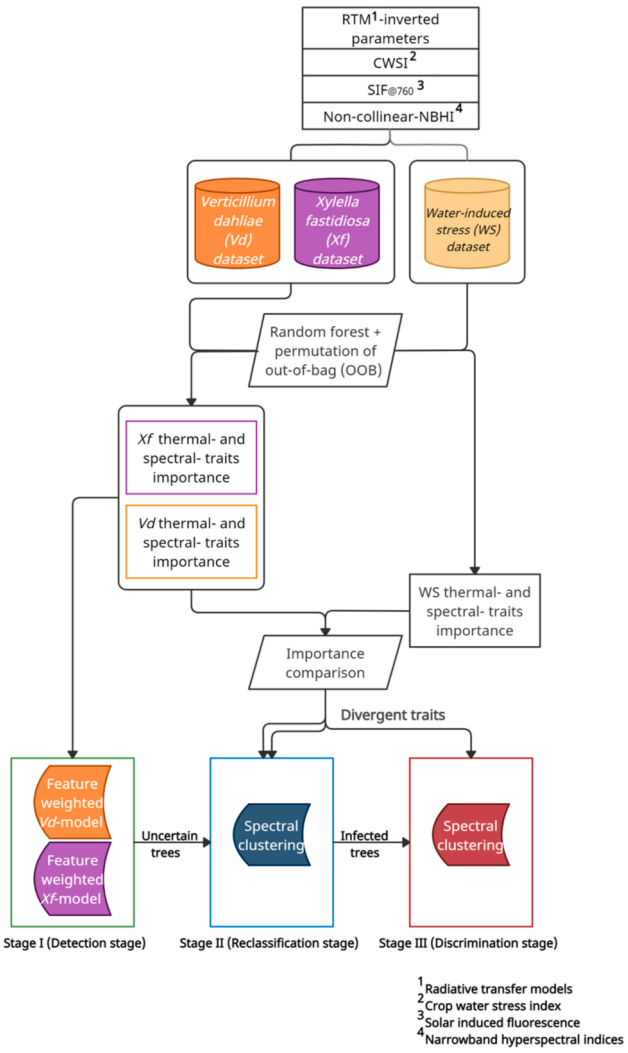

Fig. 3. Schematic representation of the three-stage classification methodology to discriminate between *Verticillium dahliae* (*Vd*) and *Xylella fastidiosa* (*Xf*) infections in olive trees.

structural properties, including leaf area index (LAI) and the leaf inclination distribution function (LIDF), were obtained from the 4SAILH model. The parameters were estimated by PRO4SAIL model inversion (Xu et al., 2019) using support vector machine (SVM) algorithms with a look-up table with 100,000 simulations. SVM models were built in MATLAB (MATLAB; Statistics and Machine Learning toolbox and Deep Learning toolbox; Mathworks Inc., Natick, MA, USA). SVM models were trained in parallel (MATLAB parallel computing toolbox), using a radial basis function and optimizing the hyperparameters during training for each of the variables described above. The inversions were performed using the simulated reflectance as inputs of the SVM models retrieving each plant trait as output. The simulations of the RTM were generated by sampling each parameter from a uniform distribution within the ranges shown in Table 1.

2.3. Statistical analyses and machine learning algorithms to discriminate between Vd- and Xf-infected olive trees

We first used analysis of variance (ANOVA) to assess whether simulated plant traits, spectral indices, and thermal traits could discriminate between asymptomatic and symptomatic trees for each disease. The ANOVA test was applied to each disease dataset, using each remote sensing trait as a response variable. Indicators were considered to significantly differentiate infected trees from non-infected trees if *p*-values were less than 0.05.

A three-stage approach was followed to discriminate *Vd*-asymptomatic (DS = 0) from *Vd*-symptomatic (DS ≥ 1) and *Xf*-asymptomatic (DS = 0) from *Xf*-symptomatic (DS ≥ 1) olive trees (Fig. 3). The three-stage process consisted of i) detection, ii) reclassification, and iii) discrimination. Predictors used in analysis included i) plant traits obtained by inverting the PRO4SAIL RTM, ii) $SIF_{@760}$, iii) CWSI, and iv) NBHIs. RTM-based traits, $SIF_{@760}$, and CWSI were included as model predictors since they directly relate to plant physiological status. Hyperspectral indices (the NBHI pool) were included after a dimension reduction step based on a Variance Inflation Factor (VIF) analysis (James et al., 2013). NBHIs with a (VIF) > 5 were considered collinear and were not included in later models (Akinwande et al., 2015). This step was important because the NBHIs include bands that are spectrally adjacent and thus redundant (Bhardwaj and Patra, 2018). The reduced, non-collinear set of NBHIs contained only those indices that improved the classification process.

(i) *Detection stage (Stage I)*. Random forests (RFs; Breiman, 2001) were trained on each disease dataset using a random split of 70% of the data, with balanced symptomatic/asymptomatic observations (TR). Models were evaluated using the remaining 30% of data (testing dataset, TS). Models were fit using all the variables at each node split and $m = 500$ trees per model. This process was iterated 100 times for each dataset (*Vf*, *Vd*, and abiotic stress), and each predictor's importance was calculated by the permutation of the out-of-bag (OOB) method (Thomas et al., 2021). Next, a feature-weighted random forest (FWRF) classification model was built for each disease using the feature importance scores for each plant trait from the initial permutation step. The FWRF algorithm was implemented in R based on the methodology proposed by Liu and Zhao (2017). The method samples distinct features in proportion to their variable importance scores and thus reduces but does not completely ignore less informative variables' contributions (Chen and Hao, 2017).

Overall, RF algorithms require the definition of two main parameters to create a classification model. These parameters include the number of trees (*m*) and the number of variables used at each node (*n*). Different studies have set the *m* parameter at its default value of $m = 500$ (Collins et al., 2020; Huo et al., 2021; Ghosh et al., 2014), since it has been suggested that *m* has a negligible effect on the accuracy compared with *n* (Belgiu and Drăguţ, 2016). Normally, for classification purposes, the number of variables randomly selected at each split in the tree building process is set as the square root of the total number of features. However, it has been suggested that setting *n* to the total number of features can give better results but can increase the computational cost (Belgiu and Drăguţ, 2016; Ghosh et al., 2014). As in our study, the total number of variables was reduced by the VIF analysis, and we included the importance of each variable into the models. All the variables at each node split were included, leaving *n* as the total number of features for each model and *m* as default ($m = 500$). Separate feature-weighted models were obtained for *Xf* and *Vd* infection datasets, each with different input feature weights.

(ii) *Reclassification stage (Stage II)*. Using the FWRF models obtained in Stage I, classification probabilities were obtained for each tree (Malley et al., 2012). Those trees that showed a probability exceeding the mean (μ) ± standard deviation (σ) limits of the probability distribution of all the trees classified were considered uncertain and were included in the clustering process described below. The traits selected for inclusion in the reclassification stage were identified as being important for predicting disease status but not for predicting abiotic stress. Only traits with disease-model importance scores at least double those of the

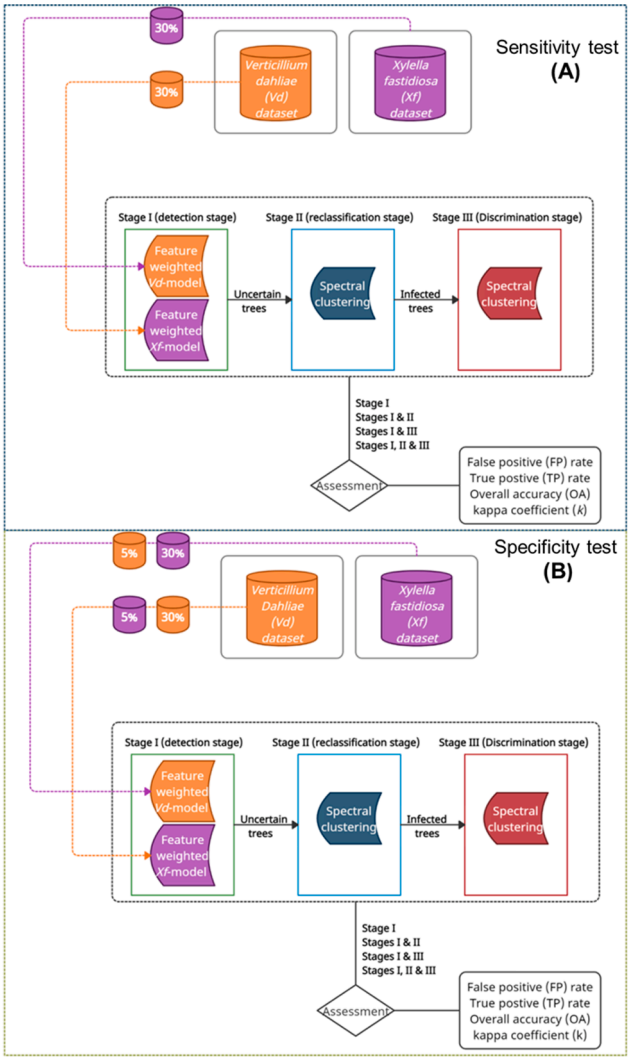

Fig. 4. Schematic representation of the sensitivity (A) and specificity (B) tests for discriminating between *Verticillium dahliae (Vd)* and *Xylella fastidiosa* (*Xf*) infections in olive trees.

corresponding importance score in the water-stress model were included as predictors in this stage.

(iii) *Discrimination stage (Stage III)*. All trees classified as infected in the previous two stages were included in the final stage. Traits used in Stage II were not considered for Stage III. The remaining plant traits whose importance at least doubled when detecting *Vd* compared with detecting *Xf* were used for the discrimination of *Vd* infections in this stage and vice versa.

For Stages II and III, a spectral clustering algorithm (Liu et al., 2013) performed in R (kernlab package, Karatzoglou et al., 2004) based on graph theory was used (Jia et al., 2014) to either discriminate diseased plants from merely stressed plants (Stage II) or discriminate one disease from the other (Stage III). This method essentially transforms the clustering problem into a graph-partitioning problem (Von Luxburg, 2007) and has been demonstrated to have superior efficiency compared to conventional clustering algorithms (Ng et al., 2001). The method can

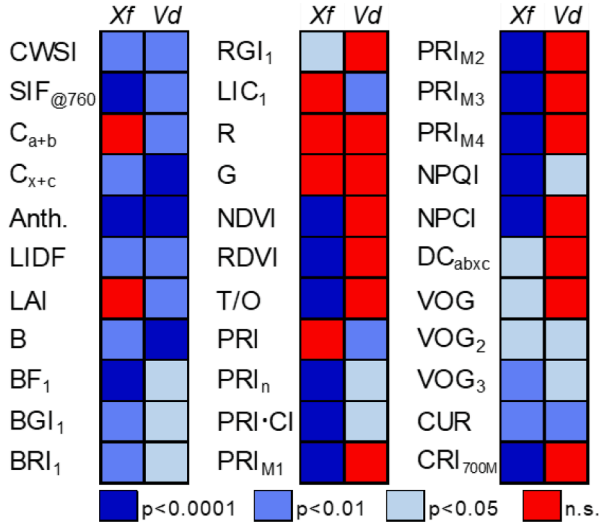

Fig. 6. ANOVA results for thermal- and spectral-derived plant traits compared between asymptomatic and symptomatic olive trees infected by either *Verticillium dahliae* or *Xylella fastidiosa*. *p* corresponds to the *p*-value obtained from the ANOVA analysis; n.s.: non-significant differences at $p \geq 0.05$.

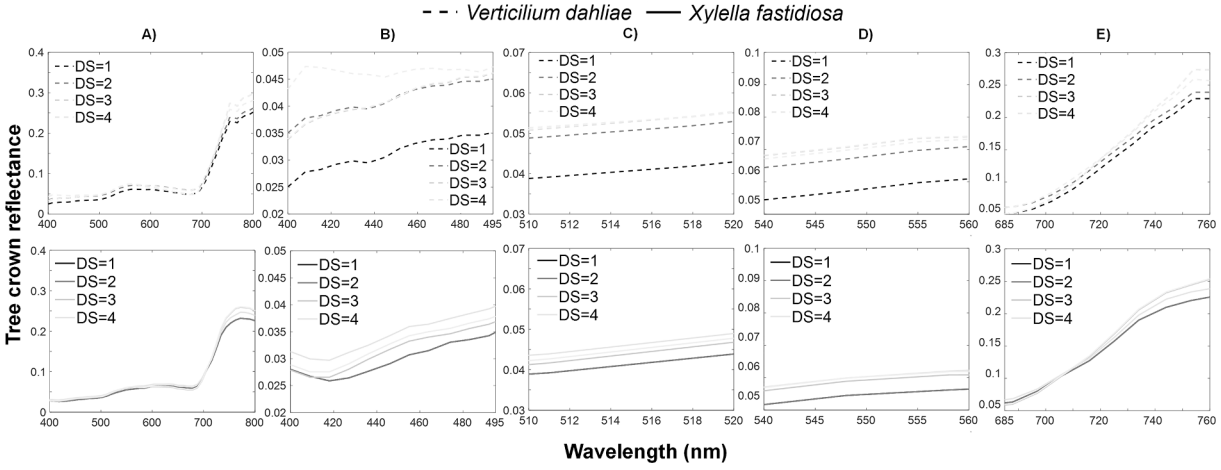

Fig. 5. Pure tree-crown airborne reflectance values for olive trees at different symptom development stages for infection by *Verticillium dahliae* (1,878 and 5,040 olive trees infected with *Vd* were assessed during 2011 and 2013, respectively) and *Xylella fastidiosa* (7,296 olive trees infected with *Xf* were assessed during 2017 and 2018). Infection statuses ranged from mild (DS = 1) to severe (DS = 4). (A) VNIR region (400–800 nm), (B) 400–495 nm (blue region of the spectrum), (C) 510–510 nm, (D) 540–560 nm, and (E) 685–769 nm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

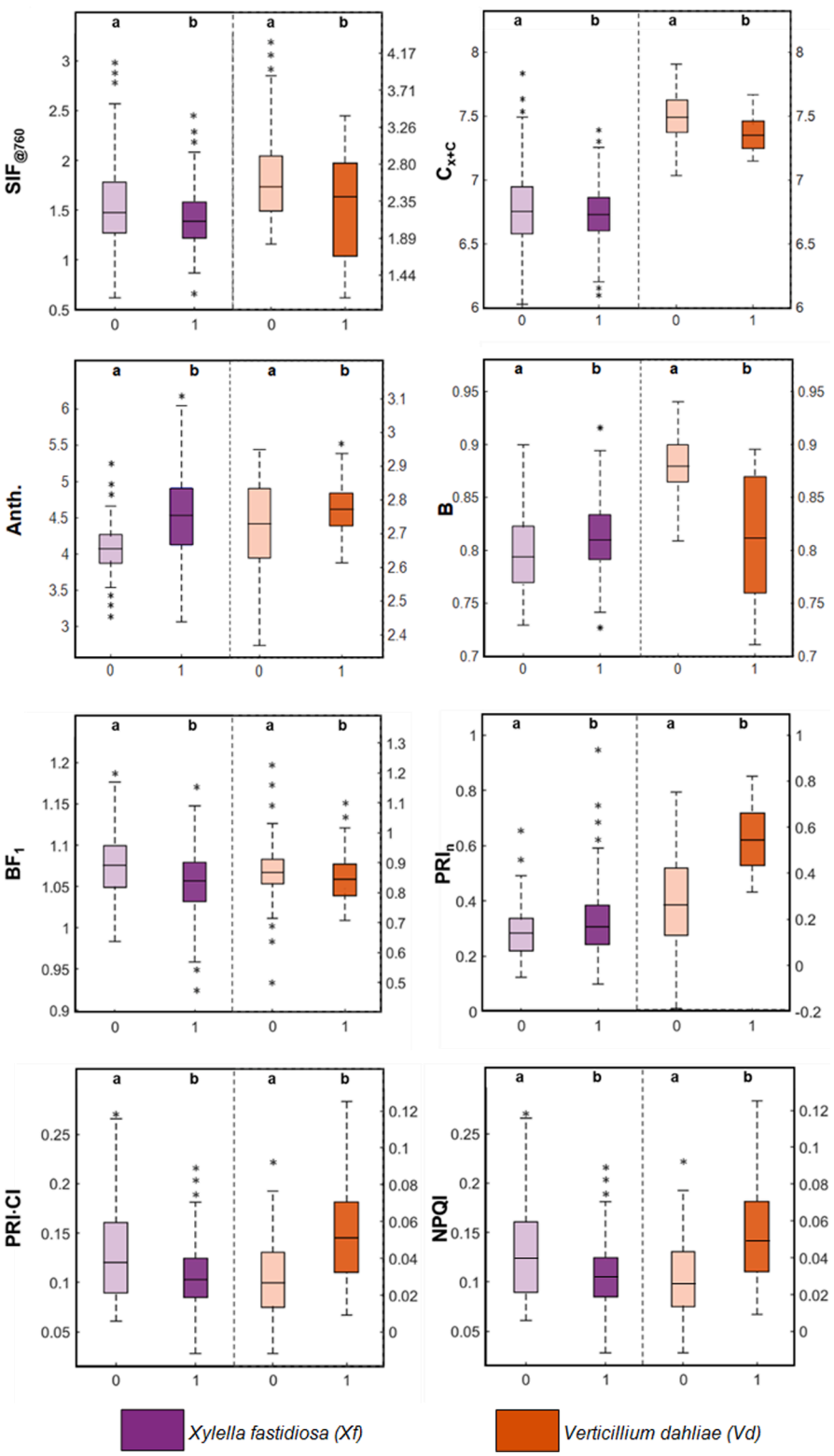

Fig. 7. Comparison of asymptomatic (0) and symptomatic (1) trees with either *Verticillium dahliae* or *Xylella fastidiosa* infections. Groups with the same infection but different letters (*a* and *b*) are significantly different according to ANOVA ($p < 0.05$). The black horizontal line in each box represents the median, and the top and bottom lines represent the 75th and 25th quartiles, respectively. The whiskers are the upper and lower limits based on the interquartile ranges (IQRs; $Q \pm 1.5 \times IQR$). The outliers (asterisks) are the values out of the upper and lower limits.

recognize clusters on unusual shapes since it performs the clustering process in a projected space in which the transformation matrix is calculated by the eigenvectors of the Gaussian similarity matrix (Tan et al., 2021). Spectral clustering has been used in remote sensing studies for classifying hyperspectral, high-resolution, and synthetic aperture radar data (Tasdemir et al., 2015; Zhao et al., 2019; Xia et al., 2015; Zhang et al., 2008). In this study, the method used to compute the affinity matrix was a Radial Basis kernel with two classes to be identified (asymptomatic versus symptomatic trees). Parameters were obtained by a heuristic procedure at every point in the dataset.

To evaluate the accuracy of the detection and discrimination between infections, specificity and sensitivity analyses were performed. As described by Trevethan (2017), specificity is understood as correctly identifying the true negatives (TN) obtained in the classification while avoiding false positives (FP). By contrast, sensitivity is understood as the probability of correctly identifying true positives (TP) while avoiding

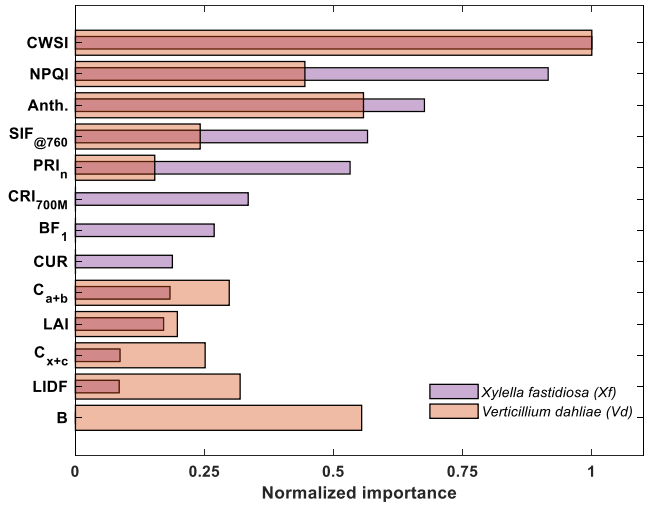

Fig. 8. Normalized variable importance scores for detecting *Verticillium dahliae*– and *Xylella fastidiosa*–infected olive trees. Scores were obtained by permutation of the out-of-bag predictor method using remotely sensed thermal and spectral traits.

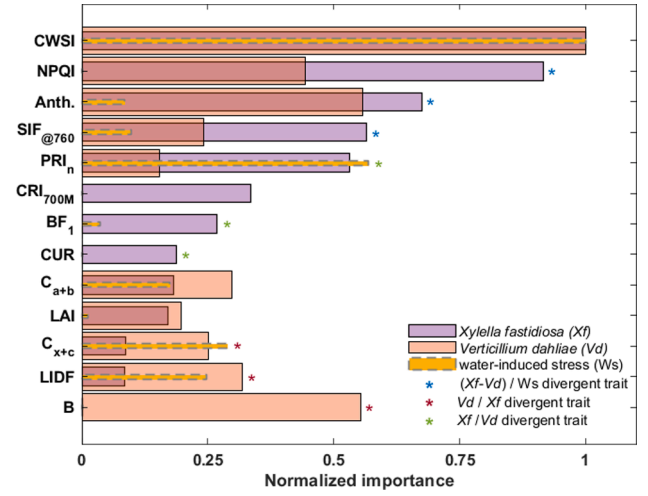

Fig. 9. Normalized importance scores for detecting *Verticillium dahliae*–infected, *Xylella fastidiosa*–infected, and water-stressed olive trees. Scores were generated for hyperspectral and thermal traits obtained by the permutation of the out-of-bag predictor method.

false negatives (FN). To assess model sensitivity, each infection model was applied to its counterpart dataset (*Xf* model on *Vd* data and vice versa). Validation datasets contained a random sample of 30% of the data, balanced between symptomatic and asymptomatic trees (Fig. 4A). A similar scheme was used to assess specificity, except that 5% of each test dataset also contained symptomatic trees of the model's target disease (*Xf* model validated on a *Vd* dataset with 5% *Xf* symptomatic trees and vice versa) (Fig. 4B). To assess any improvement in accuracy due to the three-stage approach, specificity and sensitivity tests were performed after combinations of stages, including i) Stage I; ii) Stages I and II; iii) Stages I and III; and iv) Stages I, II, and III. False-positive (FP) and true-negative (TN) rates, as well as OA and κ, were obtained and compared at each consecutive stage. In addition, errors of omission and commission were calculated considering the FP, FN, and TP rates, as shown in Anders et al. (2021).

3. Results and discussion

Visual assessments in the field confirmed that *Vd* and *Xf* infections of olive trees are difficult to distinguish due to their similarity in symptoms (Fig. 1). This is consistent with previous studies that describe similar visual symptoms for both infections (Carlucci et al., 2013). Differences in pure tree-crown reflectance between asymptomatic (DS = 0) and symptomatic (DS ≥ 1) trees were evident for both types of infections, particularly in the blue spectral region (400–495 nm) and within the ranges 510–520 nm and 540–560 nm, as well as in the red edge region at 685–760 nm (Fig. 5).

The blue region is associated with indices previously used to identify infections, including the blue ratio index (B) for *Vd* detection (Calderón et al., 2013) and the $BF_{1\text{-}5}$ and NPQI for *Xf*-induced symptom detection (Zarco-Tejada et al., 2018). The VNIR region is associated with plant pigment absorption by C_{x+c} (Chappelle et al., 1992, Merzlyak et al., 1999), Anth (Ustin et al., 2009), and C_{a+b} (Gitelson & Merzlyak, 1996). These pigment levels were retrieved in this study by RTM inversion, in contrast to previous work by Calderón et al. (2013; 2015), which only used derived reflectance indices for *Vd* detection.

ANOVA results assessing differences between asymptomatic and symptomatic trees for each disease are shown in Fig. 6. Traits that differed significantly ($p < 0.05$) between asymptomatic and symptomatic trees for both pathogens included CWSI, $SIF_{@760}$, Anth, C_{x+c}, LIDF, B, and the curvature (CUR) index. Plant traits that showed a statistically significant difference ($p < 0.05$) only for *Vd* infections but not for *Xf* infections were PRI (Gamon et al., 1992), C_{a+b}, LAI, and the Lichtenthaler index (LIC_1) (Lichtenthaler, 1996). Plant traits that were statistically different between healthy and diseased plants only for *Xf* infections were NDVI (Rouse et al., 1974), TCARI/OSAVI (Haboudane et al., 2002), CRI_{700M} (Gitelson et al., 2006), NPCI (Peñuelas et al., 1995), and the $PRI_{M(1-4)}$ indices (Gamon et al., 1992 & Hernández-Clemente et al., 2011) associated with leaf photosynthetic pigments.

Plant traits that showed highly statistically significant differences ($p < 0.0001$) for either *Vd* or *Xf* infection were considered as candidates for discriminating *Vd* and *Xf* infections. Distributions of remote sensing indices and RT model-inverted parameters for asymptomatic and symptomatic trees of each infection type are shown in Fig. 7.

Predictor importance scores from RFs (Stage I) (Fig. 8) were consistent with those reported by Zarco-Tejada et al. (2018) and Poblete et al. (2020) in detecting *Xf* infections. Zarco-Tejada et al. (2018) found that NPQI, CWSI, Anth, C_{x+c}, and $SIF_{@760}$ contributed most to discriminating between asymptomatic versus symptomatic trees, while this study found that CWSI was most important, followed by NPQI, Anth, $SIF_{@760}$, PRI_n, and CRI_{700M}. The slight differences in importance ranks for these features could be explained by the additional spectral traits used in our analyses. For example, C_{x+c} and CRI_{700M}, which are related to the carotenoid pigment content, were included in this study and found to be important for *Xf* detection. The importance of CWSI identified in this study is consistent with the results of Poblete et al. (2020), who suggested that adding CWSI to multispectral-based indices was crucial for detecting *Xf* infection. However, they also found that PRI × CI and PRI_n were at least as important. In this study, PRI_n was found to be more important than PRI × CI; thus, only PRI_n was considered.

The most important plant traits for detecting *Vd*-infected trees included CWSI, Anth, NPQI, B, and LIDF. As expected, most of these traits were similar to those selected for detecting *Xf* infection. Nevertheless, the B index (R_{490}/R_{450}) was reported as an essential indicator of *Vd* infection but not for *Xf* infection. This is consistent with the results reported by Calderón et al. (2013), in which B was also found to be sensitive to *Vd* infections in olive trees. In their study, the leaf inclination distribution function (LIDF) was not calculated and therefore not evaluated. Nevertheless, the importance of LIDF in this study is consistent with the results reported by Calderón et al. (2015), which concluded that structural indices were sensitive to the detection of *Vd*-infected trees at all stages of infection. In previous work detecting *Vd* infections,

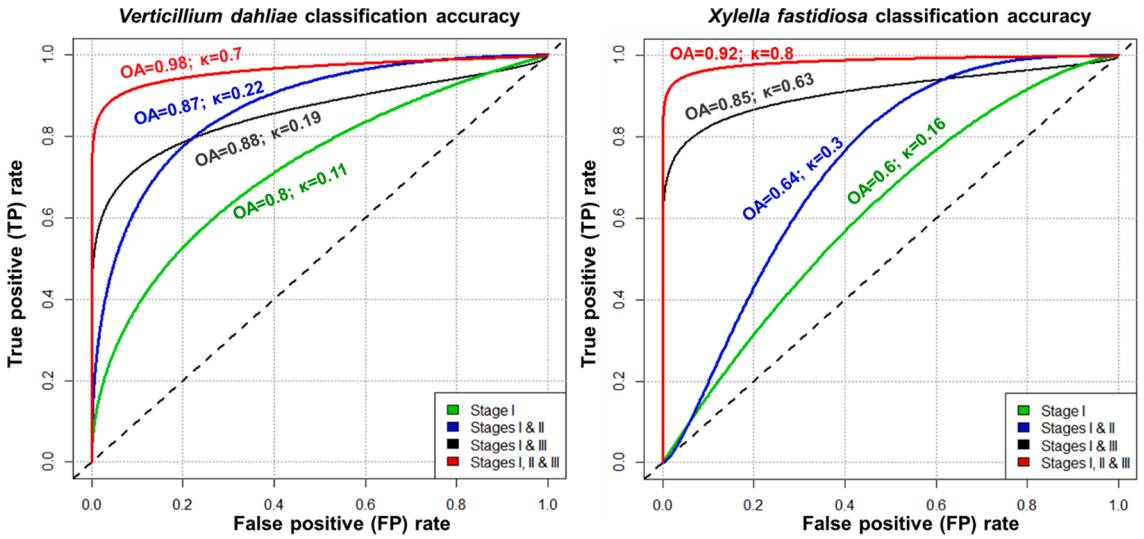

Fig. 10. Sensitivity tests for detecting *Verticillium dahliae*– and *Xylella fastidiosa*–infected olive trees. Accuracies of different combinations of methods are compared (Stage I, detection stage; Stage II, reclassification stage; and Stage III, discrimination stage). OA, overall accuracy; κ, kappa coefficient.

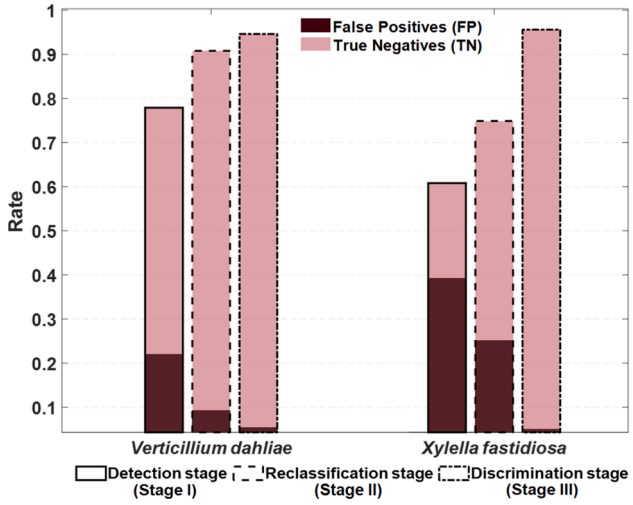

Fig. 11. Specificity test for discriminating *Verticillium dahliae*– and *Xylella fastidiosa*–infected olive trees across combinations of the detection, reclassification, and discrimination stages.

RTM inversions were not used, which might explain the decreased importance of indices such as PRI.

CWSI, $SIF_{@760}$, and PRI_n were the most important spectral plant traits suitable for detecting both diseases. However, CWSI (Calderón et al., 2013, Zarco-Tejada et al., 2018, Pineda et al., 2021), $SIF_{@760}$ (Zarco-Tejada et al., 2018, Mohammed et al., 2019), and PRI_n (Zarco-Tejada et al., 2013) have been previously shown to be highly correlated with water-induced stress in other contexts. For this reason, the reclassification stage (Stage II) was designed to reduce the uncertainty caused by traits affected by both sources of stress, i.e., *Vd*- and *Xf*-induced stress in addition to the water stress present in the orchards under study. Traits most important for detecting non-pathogenic water stress included CWSI, PRI_n, and C_{x+c} (Fig. 9). NPQI was not affected by water stress but was one of the most important indicators for *Xf* and *Vd* infection. Anth and $SIF_{@760}$ were important for detecting water-stressed trees but less important for detecting *Xf* and *Vd* infection. Due to the somewhat orthogonal distribution of importance scores among the infection and the water-stress models, the indicators NPQI, Anth, and $SIF_{@760}$ were ultimately selected as distinct inputs for the reclassification stage (Stage II).

Spectral plant traits used in the final discrimination stage (Stage III) included B, LIDF, and C_{x+c} for detecting *Vd* from *Xf* infections. PRI_n, BF_1, CUR, and CRI_{700M} traits were the most important for discriminating *Xf* from *Vd* infections.

The results of the sensitivity analysis are shown in Fig. 10. OAs based on Stage I (detection) alone were acceptable for detecting *Vd* and *Xf* (0.8 and 0.6, respectively), but κ coefficients were extremely low (0.11 and 0.14, respectively). Thus, neither model was able to accurately detect non-infected trees. The reclassification stage (Stage II) increased OA and κ for both diseases, with an overall increase of the TP rate. Nevertheless, the FP rate did not decrease, suggesting that there was no improvement in the ability to distinguish between pathogen infections. Analyses including both the detection (Stage I) and discrimination (Stage III) steps showed reduced FP rates compared to single alternatives. However, the TP rate did not increase compared to the results obtained with only Stage I.

Including the discrimination stage (Stage III) after the reclassification stage (Stage II) reduced the FP rate, but the TP rate did not drastically increase compared to using only Stages I and II. Therefore, the discrimination stage (Stage III) improved the ability to differentiate both diseases. When Stage II was omitted, both models generated less accurate predictions of infection status due to the confounding effects caused by water stress. When applying the three stages consecutively (Stage I + II + III), errors of omission were reduced to 2.9% in discriminating *Vd* from *Xf* infections and to 6.5% when discriminating *Xf* from *Vd* infections. Commission errors were 10% and 11% when discriminating between *Vd* and *Xf* infections, respectively.

The specificity analysis (Fig. 11) revealed that FP rates declined when applying the three stages consecutively. For the *Vd* model, the FP rate decreased from 0.22 to 0.09 and then to 0.05 when applying Stages I, II, and III consecutively. Similarly, FP rates for the *Xf* model decreased from 0.39 to 0.25 and then to 0.04 when the three stages were applied.

These results represent significant progress compared to previous studies focused on the detection of single infection symptoms with hyperspectral datasets either in controlled conditions (Rumpf et al., 2010; Tian et al., 2021; Zhou et al., 2019) or at the field scale (Zarco-Tejada et al., 2018; López-López et al., 2016). Nevertheless, few studies have attempted to differentiate pathogens that cause similar symptoms. Gold et al. (2020) used partial least squares discriminant analysis (PLS-DA) with hyperspectral data to distinguish between *Phytophthora infestans* and *Alternaria solani* fungal infections. Both of these diseases cause similar necrotic leaf symptoms in potatoes (*Solanum tuberosum*). They

suggested that under controlled conditions, the shortwave infrared (SWIR) spectrum best enabled differentiation between the infections. Fallon et al. (2020) used PLS-DA to detect oak wilt produced by the fungus *Bretziella fagacearum*. As a xylem-limited pathogen, oak wilt triggers symptoms similar to those of drought stress. Other diseases such as bur oak blight, caused by *Tubakia iowensis*, produce similar symptoms that could also be mistaken for oak wilt. They suggested that, under drought conditions, the differentiation between infected and uninfected leaves was possible using wavelengths on the near and SWIR spectra for two different oak species. They found that wavelengths in the range 820–1320 nm were able to differentiate between oak wilt and both drought and *T. iowensis*. They also observed small differences in the visible spectra when comparing oak wilt–infected leaves with drought-affected leaves. Our study conducted at tree-crown scales showed that the major differences between infected trees and water-stressed trees were most strongly observed in $SIF_{@760}$, NPQI, and Anth indicators.

Despite the results of Fallon et al. (2020) and Gold et al. (2020), successful remote sensing–based discrimination of tree pathologies from abiotic stresses that trigger similar symptoms has not been previously reported. This study successfully differentiated two different xylem pathogens that cause similar symptoms and proposed a method to reduce confounding effects associated with drought stress. Analysis was performed on a large spatial scale and under field conditions, using sensitive spectral indicators related to physiological plant parameters. We demonstrated that, by implementing a three-stage classification model, it was possible to discriminate between *Xf* and *Vd* infections in a combined dataset where both infections were co-occurring, yielding detection accuracies beyond 92%.

Further work will focus on the suitability of these methods across different regions and scales, accounting for large structural and background variability characteristics of different cultivars. As shown by Tian et al. (2021), infection detection accuracies decrease during the earlier stages of infection. Therefore, further work will focus on extending the proposed multiple-stage algorithm to detect and differentiate vascular diseases at early stages of infection, when abiotic confounding factors are relatively more influential. Improving the accuracy of disease detection across geographic contexts will facilitate the eradication of emerging and re-emerging pathogens worldwide.

4. Conclusions

Hyperspectral and thermal remote sensing indicators have been previously used to detect vascular pathogens such as *Vd* and *Xf* in olive trees. The most sensitive indicators previously identified include spectral traits from the inversion of radiative transfer models (Anth., C_{a+b}, C_{x+c}, LIDF), solar-induced fluorescence ($SIF_{@760}$), thermal-based CWSI, and specific narrow-band hyperspectral indices (NBHI) such as NPQI. However, both *Xf* and *Vd* infections induce similar disease symptoms, which can be confounded in spectral datasets and detection models. Here, using a combined *Vd* and *Xf* dataset, we demonstrate that a three-stage machine learning algorithm successfully discriminated between infections. The spectral traits that differentiated *Vd*-infected trees from those affected by *Xf* were the blue index B, the structural parameter LIDF, and the carotenoid pigment content C_{x+c}. Sensitivity tests revealed that OA increased from 80% to 98% when applying the three stages consecutively and that κ increased from 0.11 to 0.7. The traits found to discriminate *Xf* from *Vd* were the normalized PRI index PRI_n, the blue index BF_1, the fluorescence curvature reflectance-based index CUR, and the chlorophyll index CRI_{700M}, which improved the discrimination yielding OA from 60% to 92% and κ from 0.16 to 0.8 when using the proposed three-stage machine learning algorithm. The specificity obtained by the proposed method was high, reducing FP rates to 0.09 and 0.04 when using the cross-infected dataset to detect *Vd* and *Xf* infections, respectively. The three-stage machine learning algorithm can be used to differentiate olive tree infections, accurately identifying the source (*Vd* or *Xf*) despite the high similarity between the symptoms triggered by both xylem-invading pathogens.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank QuantaLab-IAS-CSIC staff members A. Vera, D. Notario, and R. Romero for laboratory assistance, and M. Morelli, L. Susca, M. Montes-Borrego, G. Leon and J.L. Trapero-Casas for their strong support during the field campaigns; D. Boscia and M. Saponari are acknowledged for the *Xf* field database. The study was partially funded by the European Union's Horizon 2020 Research and Innovation Programme through grant agreements POnTE (635646) and XF-ACTORS (727987), as well as by projects AGL2009-13105 from the Spanish Ministry of Education and Science, P08-AGR-03528 and P18-RT-4184 from the Regional Government of Andalusia and the European Social Fund, project E-RTA2017-00004-02 from "Programa Estatal de I + D + I Orientada a los Retos de la Sociedad" of Spain and FEDER, Intramural Project 201840E111 from CSIC. The views expressed are purely those of the writers and may not in any circumstance be regarded as stating an official position of the European Commission.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2021.07.014.

References

Akinwande, M.O., Dikko, H.G., Samson, A., 2015. Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis. Open J. Statist. 05 (07), 754–767.
Almeida, R.P.P., 2016. Can Apulia's olive trees be saved? Science 353 (6297), 346–348.
Anders, K., Winiwarter, L., Mara, H., Lindenbergh, R., Vos, S.E., Höfle, B., 2021. Fully automatic spatiotemporal segmentation of 3D LiDAR time series for the extraction of natural surface changes. ISPRS J. Photogramm. Remote Sens. 173, 297–308.
Barnes, J.D., Balaguer, L., Manrique, E., Elvira, S., Davison, A.W., 1992. A reappraisal of the use of DMSO for the extraction and determination of chlorophylls a and b in lichens and higher plants. Environ. Exp. Bot. 32 (2), 85–100.
Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.
Bhardwaj, K., Patra, S., 2018. An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images. ISPRS J. Photogramm. Remote Sens. 138, 139–150.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Bruno, G.L., Sermani, S., Triozzi, M., Tommasi, F., 2020. Physiological response of two olive cultivars to secondary metabolites of *Verticillium dahliae* Kleb. Plant Physiol. Biochem. 151, 292–298.
Calderón, R., Navas-Cortés, J.A., Zarco-Tejada, P.J., 2015. Early detection and quantification of Verticillium wilt in olive using hyperspectral and thermal imagery over large areas. Remote Sens. 7 (5), 5584–5610.
Calderón, R., Navas-Cortés, J.A., Lucena, C., Zarco-Tejada, P.J., 2013. High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of olive using fluorescence, temperature and narrow-band spectral indices. Remote Sens. Environ. 139, 231–245.
Carlucci, A., Lops, F., Marchi, G., Mugnai, L., Surico, G., 2013. Has *Xylella fastidiosa* "chosen" olive trees to establish in the Mediterranean basin? Phytopathologia Mediterranea 541–544.
Chappelle, E.W., Kim, M.S., McMurtrey, J.E., 1992. Ratio analysis of reflectance spectra (RARS): an algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. Remote Sens. Environ. 39 (3), 239–247.
Chen, Y., Hao, Y., 2017. A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. Expert Syst. Appl. 80, 340–355.
Collins, L., McCarthy, G., Mellor, A., Newell, G., Smith, L., 2020. Training data requirements for fire severity mapping using Landsat imagery and random forest. Remote Sens. Environ. 245, 111839. https://doi.org/10.1016/j.rse.2020.111839.
EFSA Panel on Plant Health (PLH), 2019. Update of the Scientic Opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory. EFSA J. 17 (5), 5665.
EFSA Panel on Plant Health (PLH), 2020. The global *Xylella* host plant database. EFSA J. 17 (5), 200.

Fallon, B., Yang, A., Lapadat, C., Armour, I., Juzwik, J., Montgomery, R.A., Cavender-Bares, J., 2020. Spectral differentiation of oak wilt from foliar fungal disease and drought is correlated with physiological changes. Tree Physiol. 40 (3), 377–390.
FAO, Food and Agriculture Organization of the United Nations, 2008. http://www.fao.org/corp/statistics/en/.
Féret, J.-B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204–215.
Fernández-Escobar, R., De la Rosa, R., Leon, L., Gómez, J.A., Testi, L., Orgaz, F., Gil-Ribes, J.A., Quesada-Moraga, E., Trapero, A., Msallem, M., 2013. Evolution and sustainability of the olive production systems. Options Mediterraneennes 106, 11–42.
Frem, M., Chapman, D., Fucilli, V., Choueiri, E., El Moujabber, M., La Notte, P., Nigro, F., 2020. *Xylella fastidiosa* invasion of new countries in Europe, the Middle East and North Africa: Ranking the potential exposure scenarios. NeoBiota 59, 77.
Gamon, J.A., Peñuelas, J., Field, C.B., 1992. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sens. Environ. 41 (1), 35–44.
Ghosh, A., Fassnacht, F.E., Joshi, P.K., Koch, B., 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. Int. J. Appl. Earth Obs. Geoinf. 26, 49–63.
Gitelson, A.A., Merzlyak, M.N., 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. J. Plant Physiol. 148 (3-4), 494–500.
Gitelson, A.A., Keydan, G.P., Merzlyak, M.N., 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. Geophys. Res. Lett. 33 (11) https://doi.org/10.1029/2006GL026457.
Gold, K.M., Townsend, P.A., Chlus, A., Herrmann, I., Couture, J.J., Larson, E.R., Gevens, A.J., 2020. Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. Remote Sens. 12 (2), 286.
Gramaje, D., Pérez-Serrano, V., Montes-Borrego, M., Navas-Cortés, J.A., Jiménez-Díaz, R. M., Landa, B.B., 2013. A comparison of real-time PCR protocols for the quantitative monitoring of asymptomatic olive infections by Verticillium dahliae pathotypes. Phytopathology 103, 1058–1068.
Gueymard, C.A., 2001. Parameterized transmittance model for direct beam and circumsolar spectral irradiance. Sol. Energy 71 (5), 325–346.
Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 81 (2-3), 416–426.
Harper, S.J., Ward, L.I., Clover, G.R.G., 2010. Development of LAMP and real-time PCR methods for the rapid detection of *Xylella fastidiosa* for quarantine and field applications. Phytopathology 100 (12), 1282–1288.
Hernández-Clemente, R., Navarro-Cerrillo, R.M., Suárez, L., Morales, F., Zarco-Tejada, P. J., 2011. Assessing structural effects on PRI for stress detection in conifer forests. Remote Sens. Environ. 115 (9), 2360–2375.
Hopkins, D.L., 1989. *Xylella fastidiosa*: xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. 27 (1), 271–290.
Huo, L., Persson, H.J., Lindberg, E., 2021. Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). Remote Sens. Environ. 255, 112240.
Idso, S.B., Jackson, R.D., Pinter, P.J., Reginato, R.J., Hatfield, J.L., 1981. Normalizing the stress-degree-day parameter for environmental variability. Agric. Meteorol. 24, 45–55.
James, G., Witten, D., & Hastie, T. (2013). Classification In: An introduction to Statistical Learning with Applications in R.
Jia, H., Ding, S., Xu, X., Nie, R.u., 2014. The latest research progress on spectral clustering. Neural Comput. Appl. 24 (7-8), 1477–1486.
Jiménez-Díaz, R.M., Cirulli, M., Bubici, G., del Mar Jiménez-Gasco, M., Antoniou, P.P., Tjamos, E.C., 2012. Verticillium wilt, a major threat to olive production: current status and future prospects for its management. Plant Dis. 96 (3), 304–329.
Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab-an S4 package for kernel methods in R. J. Stat. Softw. 11 (9), 1–20.
Klosterman, S.J., Atallah, Z.K., Vallad, G.E., Subbarao, K.V., 2009. Diversity, pathogenicity, and management of *Verticillium* species. Annu. Rev. Phytopathol. 47 (1), 39–62.
Lichtenthaler, H.K., 1996. Vegetation stress: an introduction to the stress concept in plants. J. Plant Physiol. 148 (1-2), 4–14.
Liu, J., Han, J., Aggarwal, C., & Reddy, C. (2013). Spectral Clustering.
Liu, Y., Zhao, H., 2017. Variable importance-weighted random forests. Quantit. Biol. 5 (4), 338–351.
López-Escudero, F.J., Mercado-Blanco, J., 2011. Verticillium wilt of olive: a case study to implement an integrated strategy to control a soil-borne pathogen. Plant Soil 344 (1-2), 1–50.
López-López, M., Calderón, R., González-Dugo, V., Zarco-Tejada, P.J., Fereres, E., 2016. Early detection and quantification of almond red leaf blotch using high-resolution hyperspectral and thermal imagery. Remote Sens. 8 (4), 276.
Mahlein, A.-K., Rumpf, T., Welke, P., Dehne, H.-W., Plümer, L., Steiner, U., Oerke, E.-C., 2013. Development of spectral indices for detecting and identifying plant diseases. Remote Sens. Environ. 128, 21–30.
Malley, J.D., Kruppa, J., Dasgupta, A., Malley, K.G., Ziegler, A., 2012. Probability machines: consistent probability estimation using nonparametric learning machines. Methods Inf. Med. 51 (01), 74–81.
Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L.R., Davis, C.E., Dandekar, A.M., 2015. Advanced methods of plant disease detection. A review. Agronomy Sustain. Develop. 35 (1), 1–25.
Mercado-Blanco, J., Rodríguez-Jurado, D., Parrilla-Araujo, S., Jiménez-Díaz, R.M., 2003. Simultaneous detection of the defoliating and nondefoliating *Verticillium dahliae* pathotypes in infected olive plants by duplex, nested polymerase chain reaction. Plant Dis. 87 (12), 1487–1494.
Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Rakitin, V.YU., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiol. Plant. 106 (1), 135–141.
Mohammed, G.H., Berry, J.A., Cogliati, S., Colombo, R., Damm, A., Frankenberg, C., Gastellu-Etchegorry, J.-P., Goulas, Y., Guanter, L., Joiner, J., Malenovský, Z., Meroni, M., Middleton, E.M., Miller, J.R., Moreno, J., Moya, I., Nedbal, L., Pérez-Priego, O., Rascher, U., Zarco-Tejada, P.J., 2019. Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress. Remote Sens. Environ. 231, 111177.
Navas-Cortés, J.A., Landa, B.B., Mercado-Blanco, J., Trapero-Casas, J.L., Rodríguez-Jurado, D., Jiménez-Díaz, R.M., 2008. Spatiotemporal analysis of spread of infections by *Verticillium dahliae* pathotypes within a high tree density olive orchard in southern Spain. Phytopathology 98 (2), 167–180.
Ng, A.Y., Jordan, M.I., Weiss, Y., 2001. On spectral clustering: Analysis and an algorithm. In: Proceedings of NIPS, Vancouver, BC, Canada, 3–8 December 2001, pp. 849–856.
Niblack, W., 1986. An introduction to image processing, pp. 115–116.
Peñuelas, J., Filella, I., Lloret, P., Mun Oz, F., Vilajeliu, M., 1995. Reflectance assessment of mite effects on apple trees. Int. J. Remote Sens. 16 (14), 2727–2733.
Pineda, M., Barón, M., Pérez-Bueno, M.L., 2021. Thermal Imaging for Plant Stress Detection and Phenotyping. Remote Sensing 13 (1), 68.
Plascyk, J.A., 1975. The MK II Fraunhofer line discriminator (FLD-II) for airborne and orbital remote sensing of solar-stimulated luminescence. Opt. Eng. 14 (4), 144339.
Poblete, T., Camino, C., Beck, P.S.A., Hornero, A., Kattenborn, T., Saponari, M., Boscia, D., Navas-Cortes, J.A., Zarco-Tejada, P.J., 2020. Detection of *Xylella fastidiosa* infection symptoms with airborne multispectral and thermal imagery: assessing bandset reduction performance from hyperspectral analysis. ISPRS J. Photogramm. Remote Sens. 162, 27–40.
Rodríguez-Cohard, J.C., Sánchez-Martínez, J.D., Garrido-Almonacid, A., 2020. Strategic responses of the European olive-growing territories to the challenge of globalization. European Planning Studies 28 (11), 2261–2283.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA special publication 351 (1974), 309.
Rumpf, T., Mahlein, A.-K., Steiner, U., Oerke, E.-C., Dehne, H.-W., Plümer, L., 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. Comput. Electron. Agric. 74 (1), 91–99.
Saponari, M., Boscia, D., Nigro, F., Martelli, G.P., 2013. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J. Plant Pathol. 95 (3), 668.
Sauvola, J., Pietikäinen, M., 2000. Adaptive document image binarization. Pattern Recogn. 33 (2), 225–236.
Schneider, K., van der Werf, W., Cendoya, M., Mourits, M., Navas-Cortés, J.A., Vicent, A., Oude Lansink, A., 2020. Impact of *Xylella fastidiosa* subspecies *pauca* in European olives. Proc. Natl. Acad. Sci. USA 117 (17), 9250–9259.
Sicard, A., Zeilinger, A.R., Vanhove, M., Schartel, T.E., Beal, D.J., Daugherty, M.P., Almeida, R.P.P., 2018. *Xylella fastidiosa*: insights into an emerging plant pathogen. Annu. Rev. Phytopathol. 56 (1), 181–202.
Tan, W., Sun, B., Xiao, C., Huang, P., Xu, W., Yang, W., 2021. A Novel Unsupervised Classification Method for Sandy Land Using Fully Polarimetric SAR Data. Remote Sensing 13 (3), 355.
Tasdemir, K., Moazzen, Y., Yildirim, I., 2015. An approximate spectral clustering ensemble for high spatial resolution remote-sensing images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (5), 1996–2004.
Thomas, V. A., Wynne, R. H., Kauffman, J., McCurdy, W., Brooks, E. B., Thomas, R. Q., & Rakestraw, J. Mapping thins to identify active forest management in southern pine plantations using Landsat time series stacks. Remote Sens. Environ. 252, 112127.
Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., Zhu, Y., Cao, W., Cheng, T., 2021. Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. Remote Sens. Environ. 257, 112350. https://doi.org/10.1016/j.rse.2021.112350.
Torres, M., Pierantozzi, P., Searles, P., Rousseaux, M.C., García-Inza, G., Miserere, A., Bodoira, R., Contreras, C., Maestri, D., 2017. Olive Cultivation in the Southern Hemisphere: Flowering, Water Requirements and Oil Quality Responses to New Crop Environments. Front. Plant Sci. 8.
Trevethan, R., 2017. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. Front. Public Health 5, 307.
Tsror, L., 2011. Epidemiology and control of Verticillium wilt on olive. Israel J. Plant Sci. 59 (1), 59–69.
Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A., Zarco-Tejada, P., 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sens. Environ. 113, S67–S77.
Verhoef, W., Jia, L.i., Xiao, Q., Su, Z., 2007. Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. IEEE Trans. Geosci. Remote Sens. 45 (6), 1808–1822.
von Luxburg, U., 2007. A tutorial on spectral clustering. Statist. Comput. 17 (4), 395–416.
Wong, F., Cooksey, D. A., Costa, H. S., & Hernandez, R., 2003. Documentation and characterization of Xylella fastidiosa strains in landscape hosts. In: Symposium Proceedings.

Xia, G.S., Wang, Z., Xiong, C., Zhang, L., 2015. Accurate annotation of remote sensing images via active spectral clustering with little expert knowledge. Remote Sensing 7 (11), 15014–15045.

Xu, X.Q., Lu, J.S., Zhang, N., Yang, T.C., He, J.Y., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2019. Inversion of rice canopy chlorophyll content and leaf area index based on coupling of radiative transfer and Bayesian network models. ISPRS J. Photogramm. Remote Sens. 150, 185–196.

Zarco-Tejada, P.J., González-Dugo, V., Berni, J.A.J., 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. Remote Sens. Environ. 117, 322–337.

Zarco-Tejada, P.J., González-Dugo, V., Williams, L.E., Suárez, L., Berni, J.A.J., Goldhamer, D., Fereres, E., 2013. A PRI-based water stress index combining structural and chlorophyll effects: Assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. Remote Sens. Environ. 138, 38–50.

Zarco-Tejada, P.J., Berjon, A., Lopez-Lozano, R., Miller, J.R., Martin, P., Cachorro, V., Gonzalez, M., De Frutos, A., 2005. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. Remote Sens. Environ. 99 (3), 271–287.

Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernández-Clemente, R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V., North, P.R.J., Landa, B.B., Boscia, D., Saponari, M., Navas-Cortes, J.A., 2018. Previsual symptoms of *Xylella fastidiosa* infection revealed in spectral plant-trait alterations. Nat. Plants 4 (7), 432–439.

Zhang, X., Jiao, L., Liu, F., Bo, L., Gong, M., 2008. Spectral clustering ensemble applied to SAR image segmentation. IEEE Trans. Geosci. Remote Sens. 46 (7), 2126–2136.

Zhao, Y., Yuan, Y., Wang, Q., 2019. Fast spectral clustering for unsupervised hyperspectral image classification. Remote Sensing 11 (4), 399.

Zhou, R., Jin, J., Li, Q., Su, Z., Yu, X., Tang, Y., Luo, S., He, Y., Li, X., 2019. Early detection of magnaporthe oryzae-infected barley leaves and lesion visualization based on hyperspectral imaging. Front. Plant Sci. 9, 1962.