forests

MDPI

Article

Preliminary Machine Learning-Based Classification of Ink Disease in Chestnut Orchards Using High-Resolution Multispectral Imagery from Unmanned Aerial Vehicles: A Comparison of Vegetation Indices and Classifiers

Lorenzo Arcidiaco [1], Roberto Danti [2], Manuela Corongiu [3], Giovanni Emiliani [2,*], Arcangela Frascella [2], Antonietta Mello [4], Laura Bonora [1], Sara Barberini [2], David Pellegrini [1], Nicola Sabatini [1] and Gianni Della Rocca [2]

[1] Institute of BioEconomy, National Research Council (CNR-IBE), Via Madonna Del Piano 10, 50019 Sesto Fiorentino, Italy; lorenzo.arcidiaco@cnr.it (L.A.); laura.bonora@cnr.it (L.B.); david.pellegrini@cnr.it (D.P.); nicola.sabatini@cnr.it (N.S.)
[2] Institute for Sustainable Plant Protection (IPSP), Via Madonna Del Piano 10, 50019 Sesto Fiorentino, Italy; roberto.danti@cnr.it (R.D.); arcangela.frascella@cnr.it (A.F.); sara.barberini@cnr.it (S.B.); gianni.dellarocca@cnr.it (G.D.R.)
[3] Environmental Monitoring and Modeling Laboratory for the Sustainable Development (Lamma Consortium), c/o CNR Research Area—Building D, Via Madonna Del Piano 10, 50019 Sesto Fiorentino, Italy; corongiu@lamma.toscana.it
[4] Institute for Sustainable Plant Protection (IPSP), Viale Mattioli 25, 10125 Turin, Italy; antonietta.mello@ipsp.cnr.it
* Correspondence: giovanni.emiliani@cnr.it

Abstract: Ink disease, primarily caused by the pathogen Phytophthora xcambivora, significantly threatens the health and productivity of sweet chestnut (*Castanea sativa* Mill.) orchards, highlighting the need for accurate detection methods. This study investigates the efficacy of machine learning (ML) classifiers combined with high-resolution multispectral imagery acquired via unmanned aerial vehicles (UAVs) to assess chestnut tree health at a site in Tuscany, Italy. Three machine learning algorithms—support vector machines (SVMs), Gaussian Naive Bayes (GNB), and logistic regression (Log)—were evaluated against eight vegetation indices (VIs), including NDVI, GnDVI, and RdNDVI, to classify chestnut tree crowns as either symptomatic or asymptomatic. High-resolution multispectral images were processed to derive vegetation indices that effectively captured subtle spectral variations indicative of disease presence. Ground-truthing involved visual tree health assessments performed by expert forest pathologists, subsequently validated through leaf area index (LAI) measurements. Correlation analysis confirmed significant associations between LAI and most VIs, supporting LAI as a robust physiological metric for validating visual health assessments. GnDVI and RdNDVI combined with SVM and GNB classifiers achieved the highest classification accuracy (95.2%), demonstrating their superior sensitivity in discriminating symptomatic from asymptomatic trees. Indices such as MCARI and SAVI showed limited discriminative power, underscoring the importance of selecting appropriate VIs that are tailored to specific disease symptoms. This study highlights the potential of integrating UAV-derived multispectral imagery and machine learning techniques, validated by LAI, as an effective approach for the detection of ink disease, enabling precision forestry practices and informed orchard management strategies.

Keywords: chestnut decline; *Phytophthora*; vegetation index; machine learning; support vector machine classifier; logistic classifier; Gaussian Naive Bayes classifier

check for updates

Academic Editors: Tadeusz Malewski, Piotr Borowik and Tomasz Oszako

Received: 21 March 2025
Revised: 22 April 2025
Accepted: 24 April 2025
Published: 28 April 2025

Citation: Arcidiaco, L.; Danti, R.; Corongiu, M.; Emiliani, G.; Frascella, A.; Mello, A.; Bonora, L.; Barberini, S.; Pellegrini, D.; Sabatini, N.; et al. Preliminary Machine Learning-Based Classification of Ink Disease in Chestnut Orchards Using High-Resolution Multispectral Imagery from Unmanned Aerial Vehicles: A Comparison of Vegetation Indices and Classifiers. *Forests* **2025**, *16*, 754. https://doi.org/10.3390/f16050754

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Sweet chestnut (*Castanea sativa* Mill.) is considered a relevant multipurpose tree in the EU as it plays key roles from historical–cultural, landscape, and productivity points of view (fruits and timber), being a prominent feature in the economy of many marginal territories [1,2]. The introduction and cultivation of chestnut trees through multipurpose monoculture in several southern and central European Union countries gave rise to what is usually referred to as the 'chestnut civilization' [3], to the extent that chestnut trees were called 'trees of bread' due to their importance on the diet of many rural populations until a few decades ago. Moreover, chestnut fruits and other extracts from the chestnut tree have considerable potential as nutraceutical foods or as food ingredients [4]. Sweet chestnut still covers more than 2.5 million ha of forest area in the EU [3], 780,000 ha of which are in Italy, distributed along the Apennine ridge from the north to the south of the peninsula and characterizing the social history of the Italian hills and mountains [5].

Ink disease is considered the most destructive disease to chestnuts, causing root and collar rot of trees in plantations and forest stands [1,6,7]. Ink disease is mainly caused by the invasive and widespread soil-borne oomycetes *Phytophthora xcambivora* (Petri) Buisman (previously known as *P. cambivora*) in Southern Europe and North America [8–12] although ten different *Phytophthora* species have been isolated in declining chestnut grooves [8,13]. Despite the disease has been known in Europe since the end of the 18th century, after decades of regression, the disease re-emerged in Italy in the 1990s, probably due to climate change causing increased warming and mild winters in addition to repeated drought periods [2,14]. *Phytophthora xcambivora* affects both large and feeder roots [15,16]. The course of the disease can be rapid, and the pathogen can kill adult trees in a single vegetative season, or it can have a slower progression. In the latter case, the impaired functioning of the roots leads to typical above-ground symptoms, such as the reduction in vegetative vigour and consequent sparse foliage, chlorosis, microphylly, wilting, and smaller fruits, followed by a progressive decline of the tree. In the advanced stage, the decline is characterized by the desiccation of the apical portion of the crown, which appears completely defoliated. Ink disease symptoms in the crown are best observed during the vegetative growing season, especially from July to the end of September [1,6,7,17,18].

Although a precise economic and ecological quantification of the impact of ink disease on chestnut production is not available, the disease undoubtedly affects plant survival in climate change scenarios, impacts the chestnut industry's production, and threatens the millennial-long culture of chestnut growing and the traditional landscape value, especially in Spain, Switzerland, and Italy [19,20]. In fact, for about 20 years, a resurgence of the disease has been reported in various areas in Southern Europe due to changes in the thermo-pluviometric regime, which have led to the weakening of the fine roots of plants and, consequently, a greater susceptibility to infection by the pathogen *Phytophthora xcambivora* [19].

The remote detection of tree crown health or damage status through its biophysical and biochemical properties of spectral reflectance is generating increasing interest in the fields of ecology and forestry [21–24]. A growing body of research has focused on tree health assessments through derived vegetation indices [25], and high-resolution imagery. obtained via UAV, has been used for years in cover vegetation analysis to derive both spectral and structural variables. The exploited canopy reflectance has aimed to assess disturbances and stressors in a timely and cost-effective way, helping the sustainable management of resources [26–28].

The recent advancements in image acquisition technology have facilitated the execution of sophisticated statistical investigations, and these technological innovations have opened up new avenues for object-based analysis, specifically at the tree crown level.

The advent of multispectral cameras, particularly those with four or more aligned spectral bands and exceptionally high spatial resolution (10 cm per pixel at 120 m AGL), has significantly enhanced the capacity for data collection at the crown level. Over the past decade, numerous novel methodologies for remote sensing analysis, along with the development of various optical sensors, have been introduced for terrestrial monitoring. Many of these techniques, encompassing both remote sensing and image analysis, have been successfully applied to determine the spatial distribution of plant pathogens [28–30], including vegetation decline due to *Phytophthoras*, especially in the agricultural field [31–36] but also in forests [37–39].

The selection of the eight vegetation indices—normalized difference vegetation index (NDVI), green normalized difference vegetation index (GnDVI), red-edge normalized difference vegetation index (RdNDVI), excess green-excess red index (ExGreenRed), enhanced vegetation index (SAVI), enhanced vegetation index (EVI), enhanced vegetation index (EVI_2), and modified chlorophyll absorption in reflectance index (MCARI)—was driven by their proven relevance in previous literature to vegetation health monitoring, stress detection, and spectral sensitivity to chlorophyll, biomass, and canopy structure. These indices were chosen based on the following criteria:

- Physiological sensitivity to stress symptoms: Given that Phytophthora spp. infection often manifests initially through subtle changes in chlorophyll content, leaf structure, and canopy density, we selected indices known for their ability to detect vegetation stress and physiological deterioration [40,41];
- Diversity of spectral characteristics and mathematical formulations: The indices span a range of spectral regions (visible, near-infrared, and red-edge) and incorporate various correction mechanisms (e.g., for soil background or atmospheric effects), allowing us to test and compare the relative performance of VIs under different spectral sensitivities;
- NDVI remains the most widely used VI for assessing general vegetation vigor and health [42]; GnDVI replaces the red band with the green band, improving sensitivity to chlorophyll content [43]; RdNDVI leverages the red-edge band, which has shown superior performance in detecting subtle physiological stress responses in plant canopies [43,44]; MCARI emphasizes chlorophyll absorption and is particularly well-suited for detecting changes in leaf pigment concentration [45]; SAVI and EVI/EVI_2 offer soil and atmospheric correction capabilities, which help minimize non-canopy signal contamination, which is particularly important when ground exposure varies due to canopy thinning [46–48]; ExGreenRed was included to evaluate the performance of simpler spectral combinations in vegetation detection.

By including a diverse suite of indices—some emphasizing structural traits (e.g., NDVI, RdNDVI), others pigment concentration (e.g., MCARI), and some correcting for environmental noise (e.g., SAVI, EVI)—we aimed to comprehensively evaluate which VI–classifier combinations best discriminate symptomatic from asymptomatic trees.

At the same time, classifiers based on machine learning algorithms (SVM, Log, GNB) or neural networks based on deep learning techniques are increasingly being employed by researchers and demonstrate a significant improvement in classification performance [31,49,50].

Pixel-based classification techniques for individual trees have already been used for a few years [51,52], and include deep learning methods like convolutional neural networks (R-CNNs) [23,53] for assessing individual tree features using different VIs [21]. A previous attempt to identify the foci of ink disease caused by *Phytophthora xcambivora* in a chestnut forest through remote sensing images was realized in the Italian Apennines [54]. Recently, Padua et al. [55] applied ML techniques to multispectral data derived from UAVs in chestnut trees in Portugal for the identification of issues of biotic and abiotic origin.

Within the framework of the LIFE MycoRestore project (LIFE 18/CCA/ES/001110), we conducted a detailed analysis based on multispectral orthophotos constructed from images acquired by UAVs, where spectral characteristics associated with plant crowns were derived and analyzed. The analysis aimed to use the most common vegetation indices to measure the health status of the vegetation itself. Three different classification algorithms were modeled, and the performance results made it possible to assess which combination of classifier–VI is best suited to discriminate the health status of plants.

Another objective of this work was to model and evaluate the results of three implemented classifiers that were applied to all VIs. This evaluation aimed to determine which classifier, with its equivalent VI, showed higher performance. Additionally, with the same classifier, the research sought to identify which VIs were most suitable for assessing the chestnut trees' health status. The selection of three classifiers (SVM, GNB, and Log) for this study was based on several considerations regarding the characteristics of the predictors and the classification objective. The SVM is a powerful classifier, particularly for binary classification problems, due to its ability to identify an optimal hyperplane that maximizes the margin between the two classes. In scenarios where predictors, such as vegetation indices, are clearly separable into two distinct classes, the SVM performs well, even when the data are not linearly separable, by leveraging non-linear kernels. Its robustness against overfitting, particularly in high-dimensional datasets, makes it an excellent choice when a clear class separation is expected [56–58].

The GNB is well-suited for data that follow a normal distribution, as is often the case with vegetation indices, which typically exhibit symmetric or bell-shaped distributions. This classifier is simple, fast to train, and particularly effective when the predictor variables are independent or can be approximated as conditionally independent given the class. Despite its independence assumption, the GNB classifier can be surprisingly robust and efficient, especially when the classes are well separated, as in the present case [59,60]. Finally, the Log classifier was chosen because it is another powerful model for binary classification that provides class membership probabilities. This model assumes a linear relationship between the predictor variables and the log-odds of the target class [61].

Classifiers were selected upon the following criteria: (a) their demonstrated effectiveness in high-dimensional, limited-sample scenarios, and (b) their interpretability in correlating spectral features with physiological tree stress. In contexts where the predictors, such as vegetation indices, allow for clear class separation, logistic regression is particularly suited to model these linear relationships. Their simplicity, interpretability (e.g., regression coefficients), and ability to handle large datasets without requiring complex computations make them ideal for binary classification problems. In general, the use of these three classifiers provides comprehensive coverage of different aspects of the problem. SVM offers strong separation capabilities even in complex scenarios, Naive Bayes provides a fast and efficient solution for Gaussian-distributed data, and logistic regression is simple and effective for linear relationships. Combining these methods allows for performance comparison and model selection based on the actual data, thereby enhancing the reliability of the classification results (SVM [56–58], GNB [58–60], and Log [61]).

Ink disease caused by *Phytophthora* spp. represents a major threat to chestnut orchards, yet early detection efforts have been hampered by its subtle, canopy-level symptoms that traditional remote sensing approaches often fail to capture until the disease has progressed to moderate or severe stages. Although various multispectral and synthetic aperture radar satellite platforms have recently been employed to map symptomatic trees, they typically rely on the more pronounced spectral differences that arise once significant foliar decline has occurred, thus overlooking earlier, less visible indicators of infection, and the discrimination among disease classes of different severity (moderate and severe damage)

is less accurate [62]. This gap underscores the need for a refined research question that explicitly addresses why existing approaches, which perform moderately well for advanced disease mapping, are insufficient for timely assessments of incipient infections. Our study aims to fill this void by evaluating high-resolution multispectral datasets in conjunction with machine learning classifiers specifically tailored to identify subtle reflectance changes linked to stress responses. Through this integrated methodology, we seek both to clarify the limitations of current remote sensing strategies and to demonstrate how targeted data acquisition combined with advanced analytical techniques can significantly improve the prompt detection of ink disease symptoms in chestnut stands.

2. Materials and Methods

2.1. Site Description, Dendrometric Measurement, and Health Evaluation of Trees

This work was carried out in an old chestnut orchard for fruit production located in Castagno d'Andrea (box boundary: ≻upper-left 4864333N-715094E; upper-right ≻4864332N-715408E; bottom-right ≻4864141N-715410E; bottom-left ≻4864141-715090N; -780 m a.s.l., orientation E-NE, Figure 1), San Godenzo Municipality in the Tuscan Apennines (Mugello area, Tuscany, Italy), which is partially affected by ink disease [11,14]. According to the pedological map of the Tuscany Region, the soil is classified as Typic Hapludalfs, fine silty, mixed, mesic (Soil Taxonomy, 2022; [63]) and Cutanic Luvisols (Classification WRB, [64]); the soil texture corresponds to loam soil. The climatology of the sites is sensitized by the Walter–Leight diagram in Figure 2, which shows that the probable freeze period is January to February and December; the average annual temperature is 6.7 °C; the mean absolute maximum temperature is 30.4 °C, and the average absolute minimum temperature is −2.3°; the average annual precipitation is 994 mm. All climate parameters were derived from the development of the ERA5-Land climate dataset [65], which is produced and distributed by the European Center for Medium-Range Weather Forecasts; Copernicus Climate Change Service (C3S: https://climate.copernicus.eu/, accessed on 15 November 2023).

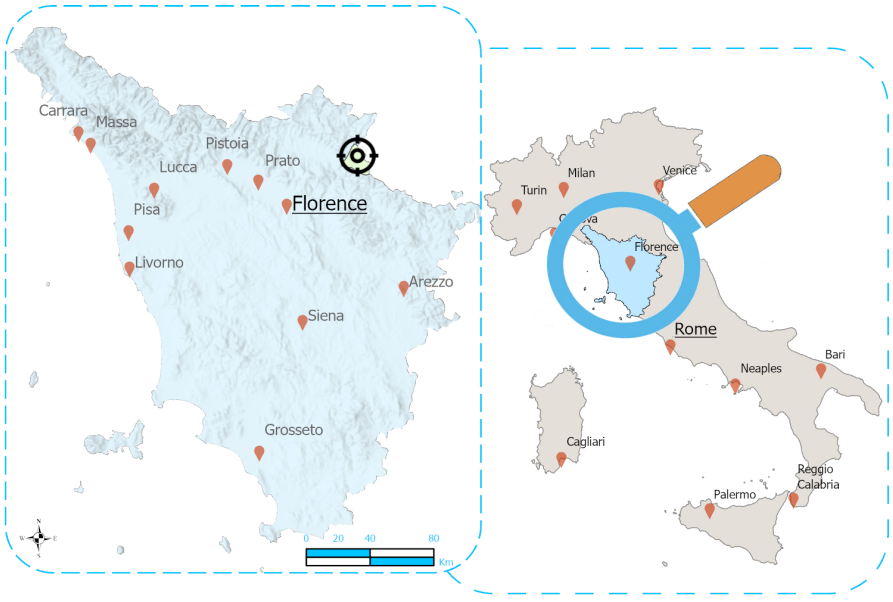


Figure 1. Area of interest: in the maps of Italy (**right** image) and Tuscany (**left** image), the location of the chestnut orchard where this study (Castagno d'Andrea, municipality of San Godenzo, Florence district) was performed is shown by a circular black pin (**left** image).

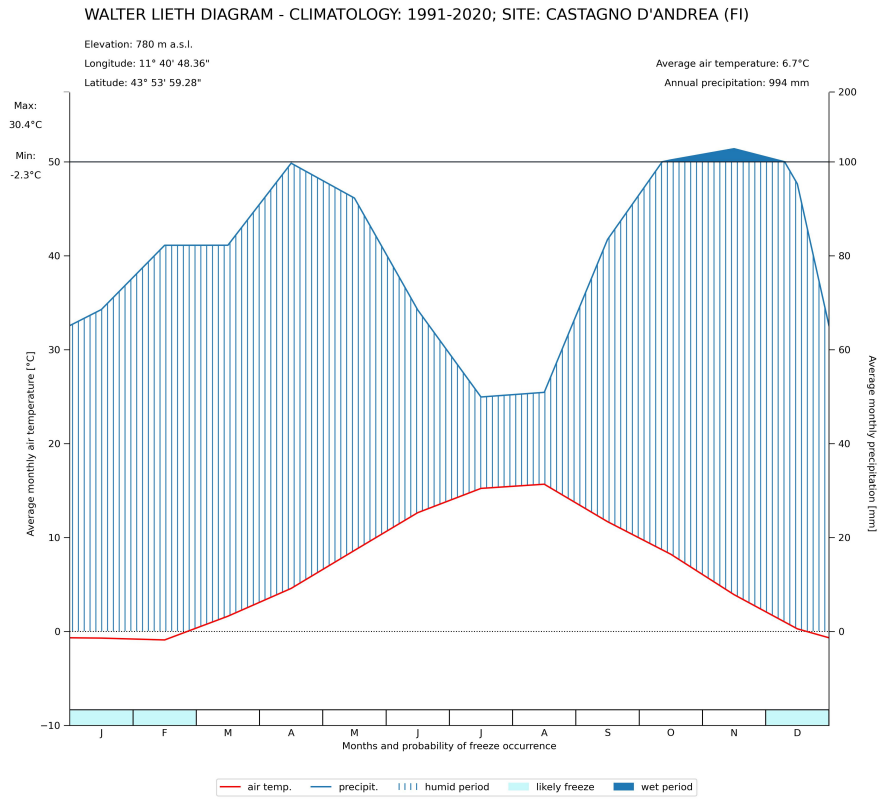


Figure 2. Walter–Lieth diagram (climatology 1991–2020) of Castagno d'Andrea (San Godenzo, Italy), where the chestnut orchard is located.

In the chestnut orchard, two areas were identified: the first, called 'asymptomatic', consisted of apparently healthy chestnut trees (29 plants), and the second, called 'symptomatic', in which trees showed different symptomatic stages attributable to ink disease (38 trees) [11,14]. As reported by Venice et al. [14] all the trees growing in both areas of the chestnut orchard were visually assessed for their health status according to Vannini et al. [16,66] and Maresi et al. [67], and some dendrometric parameters were measured (diameter at breast high, tree height, and crown width). In detail, the presence of crowns with typical symptoms of ink disease were evaluated according to Schomaker [68] and the description from the Italian National Forest Inventory (INFC, https://www.inventarioforestale.org/en/), while crown mortality was evaluated following the images from Bosshard [69].

For all trees, the absolute and precise location was acquired using a TOPCON HIPER HR Global Navigation Satellite System (GNSS) instrument [70] (Topcon Positioning Systems Inc., Tokyo, Japan) adopting reference system WGS84-UTM32; EPSG:32632. The Hiper HR GNSS can acquire and store positions with sub-metric precision in real-time kinematic (RTK) positioning using the Vanguard Technology^{TM} with universal tracking channels for multi-frequency tracking of multiple satellite constellations, such as GPS, GLONASS, BeiDou, QZSS, SBAS, and Galileo. The possibility of relying on several constellations at the same time allows from time to time to select the one that provides the best accuracy and thus greater precision in the submetric survey.

The post-processing analysis was carried out with Magnet Office^{TM} software [71] (Release 8.0) dedicated to the post-processing of GNSS data (Topcon Positioning Systems Inc., Tokyo, Japan). All trees were identified with a unique identifier number (UID) tree code.

In addition, ten chestnut trees from each of the symptomatic and asymptomatic areas, located at least 20 m apart, were randomly selected and subjected to leaf area index

(LAI) measurements (Table 1) using an LAI-2000 (Li-cor, Lincoln, NE, USA) six times under each canopy to estimate the coverage of leaves on the ground [14,72,73]. Validation of symptom etiology was also performed using the baiting method for the isolation of *Phytophthora xcambivora* from the soil, according to Erwin and Ribeiro [15], as described in Venice et al. [14].

Table 1. Leaf area index measures and dendrometric parameters of 10 visually classified symptomatic and 10 asymptomatic chestnut trees.

Plant UID	Circunf. (cm)	High (m)	Crown Width (m)	Symptomatic (S)/ Asymptomatic (A)	Crown Mortality (%)	LAI
30	328	15.2	14.3	A	0–10	8.25
105	270	14.5	12.0	A	0–10	7.58
110	255	16.9	11.5	A	0–10	8.30
130	225	12.4	9.3	A	0–10	7.58
155	300	18.3	13.0	A	0–10	9.87
210	540	19.6	15.0	A	0–10	8.56
225	270	21	15.0	A	0–10	10.36
315	245	15.7	12.0	A	0–10	6.58
965	350	12.7	13.0	A	0–10	7.74
1075	230	16.1	11.8	A	0–10	7.47
Average for A group	*297.7 cm*	*15.4 m*	*12.5 m*		*0–10*	*8.27*
25	350	16.5	10.0	S	25–50	5.91
35	255	18.4	7.2	S	50–99	2.60
50	330	13.6	10.3	S	50–99	6.30
94	290	15.8	9.5	S	25–50	5.05
95	325	10.4	10.4	S	10–25	4.55
96	200	17.6	9.9	S	25–50	5.72
97	180	16.2	8.6	S	50–99	6.35
98	160	13.4	4.0	S	50–99	3.22
99	210	10.5	9.5	S	0–10	6.99
105	200	14	8.5	S	25–50	5.91
Average for S group	*248.4 cm*	*14.7 m*	*8.6 m*		*50–80*	*5.25*

2.2. Surveyed Area and Data Acquisition

For this research, a fixed-wing and tailsitter vertical take-off and landing (VTOL) UAV WingtraOne RX1R Drone (Wingtra AG-Zurich, Zürich, Switzerland) was used to collect the aerial imagery (Figure 3a). The drone was equipped with a compact multispectral camera MicaSense RedEdge-M (MicaSense, Inc.; now AgEagle Aerial Systems Inc., Seattle, WA, USA) [74], which can acquire 5 different bands simultaneously: blue$_{475\text{nm wavelength}}$ (B_{475}), green$_{560\text{nm wavelength}}$ (G_{560}), red$_{668\text{nm wavelength}}$ (R_{668}), near infrared$_{840\text{nm wavelength}}$ (NiR_{840}), and red edge$_{717\text{nm wavelength}}$ (RE_{717}). The camera had a focal length of 5.5 mm and captured images at a resolution of 1280 × 960 (Tables 2 and 3).

Table 2. Spectral characteristics of the MicaSense RedEdge-M multispectral camera sensors.

Sensor	Central Wavelength (nm)	Filter Bandwidth [1] (nm)
Blue (band 1)	475 (B_{475})	20
Green (band 2)	560 (G_{560})	20
Red (band 3)	668 (R_{668})	10
Near-IR (band 4)	840 (NiR_{840})	40
Red-Edge (band 5)	717 (RE_{717})	10

[1] Full width at half maximum.

Table 3. Lens and imager information of MicaSense RedEdge-M multispectral camera.

MicaSense RedEdge-MX	
Pixel size	3.75 µM
Resolution	1280 × 960 (1.2 MP × 5 imagers)
Aspect ratio	04:03
Sensor size	4.8 mm × 3.6 mm
Focal length	5.5 mm
Field of view	47.2 degrees horizontal; 35.4 degrees vertical
Output bit depth GSD @ 120 m (∼400 ft)	12-bit 8 cm/pixel per band
GSD @ 60 m (∼200 ft)	4 cm/pixel per band

The flight was conducted over the chestnut orchard after the summer period on 2 October 2021, when the higher temperatures and lower rainfall typical of the summer period had exacerbated the symptoms of the pathogen on the vegetative condition of the infected trees [6,9,16].



(**a**) WingtraOne RX1R drone



(**b**) White calibration panel

Figure 3. WingtraOne drone (**a**) and white panel (**b**) used for camera calibration.

The flight was performed close to solar noon time to minimize shadows, and the above-ground levels (AGLs) of the survey were calculated taking into account all inherent parameters: camera sensor resolution, focal length, desired ground sample resolution, camera shutter speed, UAV translation, and speed, at about 110–120 m altitude with a constant forward speed of 10 m s^{-1}. The dedicated WingtraPilot software (Release 2.0—Wingtra AG-Zurich, Switzerland) was used to plan the flights, in which the user defines the area of interest, flight direction, longitudinal and lateral overlapping, and ground sample distance (GSD). The flight was carried out using the typical aerial mapping profile 'serpentine' with high percentages of superposition on the whole track and along the track to collect enough data to rebuild the entire high-resolution orchard image. The photos were taken with a forward overlap of 70% and a lateral overlap of 80%. All the 7170 aerial photos

were acquired with a GSD average of ≃10 cm. The high resolution of the orthophotos allowed a 'pixel-based approach at individual canopy level' analysis and a comparative performance of vegetation indices and machine learning classifiers for the detection of ink disease in chestnut trees.

2.3. Image Processing, Maps Production, and Extraction of VIs

Spectral reflectance for each band was calibrated and normalized using calibration images and the appropriate correction factors of a white calibration panel (Figure 3b; panel code: RP06-2114009-OB). The images were aligned by matching the tie points across all adjacent photos, using the structure from motion (SfM) photogrammetry technique [75]. Photogrammetric processing enables the achievement of advanced geographic data products, such as multispectral stacked images (orthophoto mosaics, Figure 4a,b), different VI images (Figure 4c), digital surface models (DSMs) (Figure 4d), digital terrain models (DTMs) (Figure 4e), and three-dimensional (3D) point clouds. The spatial resolution for orthophoto mosaics was 0.25 m/pixel; while for DSM and DTM, it was 0.5 m. The aerial images were processed in Agisoft PhotoScan Professional (version: 1.8 build 13111, 64 bit Agisoft LLC, St. Petersburg, Russia).

From the final orthophoto mosaics we extracted five spectral bands (blue, green, red, red-edge, near-infrared) at a spatial resolution of [0.25 m/pixel]. Radiance values were converted to surface reflectance based on calibration panel readings and manufacturer guidelines.

Using the reflectance values, 8 distinct VIs were calculated through the spatial analysis operations detailed in Table 4. The advantage of analyzing vegetation indices instead of individual spectral bands lies in the fact that the relationship between vegetation indices and the eco-physiological behavior of plants is significantly stronger and more robust compared to that of individual bands [76,77].

(**a**) Red–green–blue true color composition

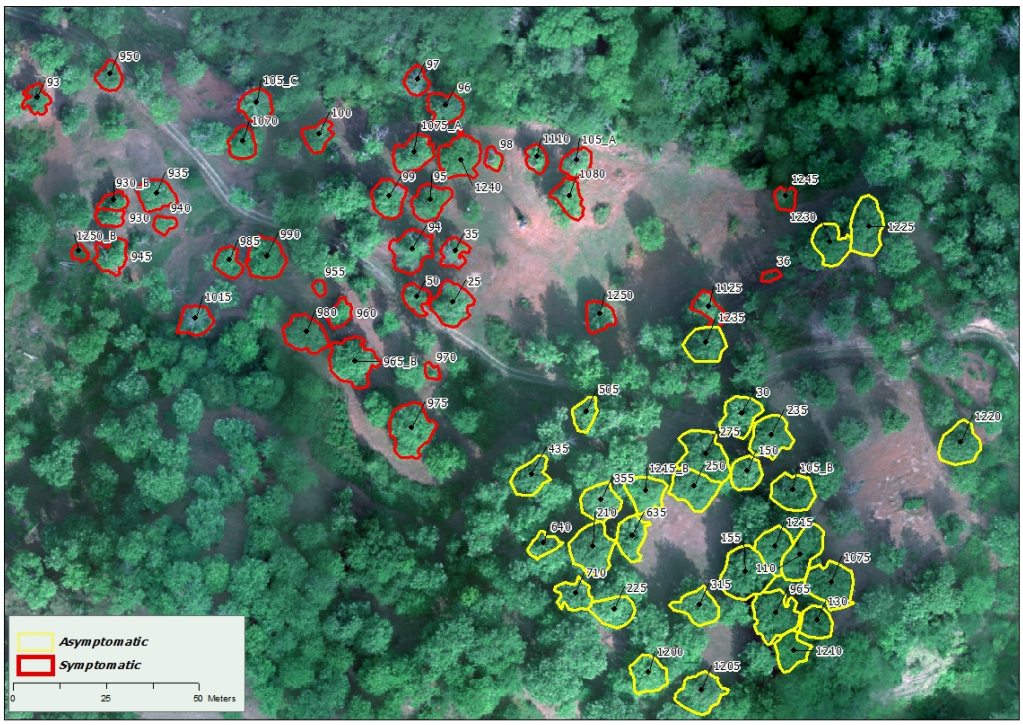
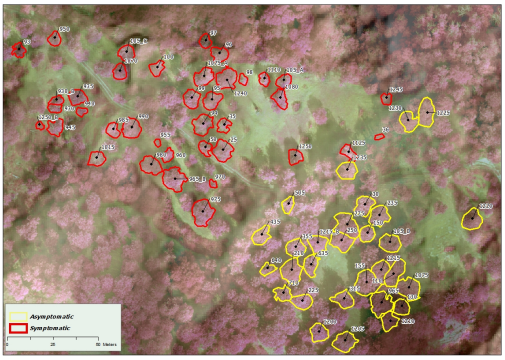
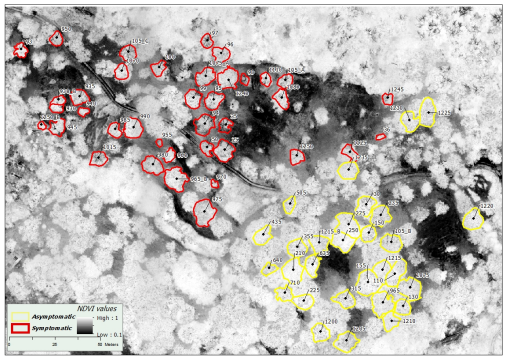


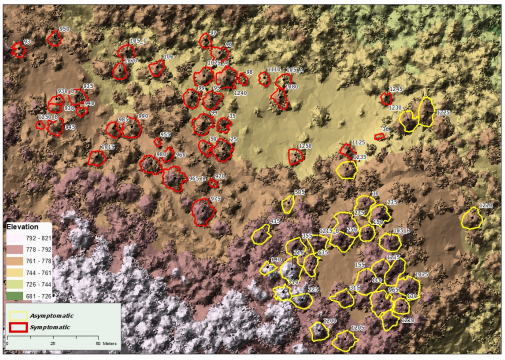
Figure 4. *Cont.*



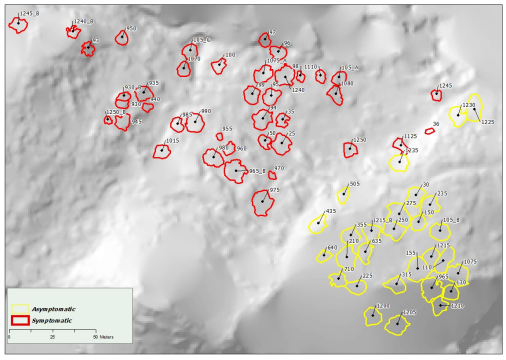
(**b**) Nir–red–green false color composition



(**c**) NDVI map.



(**d**) Digital surface model (DSM).



(**e**) Digital terrain model (DTM).

Figure 4. Significant maps of the area of interest. Red = symptomatic trees; yellow = asymptomatic trees.

Table 4. VIs used for this research and their formulas.

Index	Brief Explanation	Formula		Range	Reference
NDVI	Normalized difference vegetation index. Most commonly used to estimate vegetation health and biomass.	$\frac{NiR_{840} - R_{668}}{NiR_{840} + R_{668}}$	(1)	−1 to +1	[78]
GnDVI	Green NDVI. Uses green reflectance instead of red, often more sensitive to chlorophyll content.	$GnDVI = \frac{NiR_{840} - G_{560}}{NiR_{840} + G_{560}}$	(2)	−1 to +1	[43]
ExGreenRed	Excess green minus excess red. A color-based index (RGB) used in vegetation detection from standard cameras.	$(2 \times G_{560} - R_{668} - B_{475}) - (1.4 \times R_{668} - B_{475})$	(3)	∞ to +∞	[79]
RdNDVI	Similar to NDVI but uses the red-edge band instead of red, improving sensitivity to changes in canopy structure.	$\frac{NiR_{840} - RE_{717}}{NiR_{840} + RE_{717}}$	(4)	−1 to +1	[43]
SAVI	Soil adjusted vegetation index. Reduces soil background reflectance using a factor L.	$2.5 \times \frac{(NiR_{840} - R_{668}) \times (1 + L)}{NiR_{840} + R_{668} + L^*}$ **where L is a correction factor: L = 0.428*	(5)	−1 to +1	[46,47]

Table 4. *Cont.*

Index	Brief Explanation	Formula		Range	Reference
EVI	Enhanced vegetation index. Optimized to enhance vegetation signals in high-biomass areas.	$2.5 \times \frac{NiR_{840} - R_{668}}{NiR_{840} + C1^* \times R_{668} - C2^* \times B_{475} + 1}$ *where C1 and C2 are two correction factor: C1 = 6; C2 = 7.5	(6)	−1 to +1	[47]
EVI_2	A two-band version of EVI that removes the need for a blue band.	$2.5 \times \frac{NiR_{840} - R_{668}}{NiR_{840} + R_{668} + 1}$	(7)	0 to +2	[48]
MCARI	Modified chlorophyll absorption in reflectance index. Emphasizes chlorophyll absorption in the red region.	$\left((RE_{717} - R_{668}) - 0.2 \times (RE_{717} - G_{560})\right) \times \left(\frac{RE_{717}}{R_{668}}\right)$	(8)	0 to 1	[45]

All VI computations were performed in programming language Python 3.8 using Rasterio [80] and NumPy [81] libraries. The resulting VI image rasters were stored at the same spatial resolution and coordinate reference system as the original orthomosaic. The mean values per plant of each VI were then correlated using a linear regression with acquired LAI. The value of R^2 was calculated using the formula reported in Equation (9).

$$R^2(y, \hat{y}) = \frac{\sum_{i=0}^{N-1}((y_i - mean(y)) \times (\hat{y}_i - mean(\hat{y})))}{\sqrt{\sum_{i=0}^{N-1}(y_i - mean(y))^2} \times \sqrt{\sum_{i=0}^{N-1}(\hat{y}_i - mean(\hat{y}))^2}} \quad (9)$$

2.4. Pixel Extraction

As reported in Section 2.2, we conducted a field survey to identify and label individual chestnut trees based on visual symptoms of ink disease. The plants were grouped according to their health status: symptomatic and asymptomatic. For each VI, the distribution graph of pixels belonging to symptomatic and asymptomatic plants was created. For each tree in both asymptomatic and symptomatic areas (see Section 2.1), the crown was digitized by on-screen photointerpretation using ArcGis Pro software (release 3.2) [82], adopting an overlay raster extraction technique; for each tree crown, it was possible to extract all the pixels from all VI images. This image processing allowed us to obtain a probability density function (PDF) of all VI values. To perform raster extraction of VI values and obtain the PDF, a specific algorithm was performed in the Anaconda ecosystem [83] using the Python programming language and Rasterio, Numpy, Pandas, and Statsmodels modules.

To support the visual classification of the health status of chestnut tree crowns, linear regressions between LAI and VI average values of chestnut crowns were implemented. In addition, for four VIs (RdNDVI, NDVI, GnDVI, and EVI_2: indices whose calculation involves the use of only two bands), isocline plots were created to visualize the better-discriminating index value that can separate asymptomatic and symptomatic chestnut trees. In order to derive isocline plots of the VIs, for each crown, the average spectral reflectance value was derived from the two bands involved. Those values and the isoclines, representing the VI values at the various steps, were plotted.

2.5. Data Analysis and Modeling the Classifiers

The step of pixel extraction corresponding to each canopy allowed us to obtain arrays of all VIs for each tree, facilitating the implementation of various statistical analyses. For each VI and each plant, essential statistical parameters (mean, median, first quartile,

third quartile, minimum value, maximum value, range) were computed. Additionally, the PDFs were analyzed and their respective cloud plots were derived. All plants underwent a grouping operation based on their health status, resulting in two PDFs for each VIs: one for asymptomatic plants and one for symptomatic ones, both reported as boxen plots.

From the total set of crown-level data points, using scikit-learn module, we randomly allocated 70% to the training set and 30% to the test set. We maintained a balanced proportion of symptomatic and asymptomatic samples in both sets.

The selection of three classifiers (SVM, GNB, and Log) for this study was based on several considerations regarding the characteristics of the predictors and the classification objective. The SVM is a powerful classifier, particularly for binary classification problems, due to its ability to identify an optimal hyperplane that maximizes the margin between the two classes. In scenarios where predictors, such as vegetation indices, are clearly separable into two distinct classes, the SVM performs well, even when the data are not linearly separable, by leveraging non-linear kernels. Its robustness against overfitting, particularly in high-dimensional datasets, makes it an excellent choice when a clear class separation is expected [56–58].

The GNB is well-suited for data that follow a normal distribution, as is often the case with vegetation indices, which typically exhibit symmetric or bell-shaped distributions. This classifier is simple, fast to train, and particularly effective when the predictor variables are independent or can be approximated as conditionally independent, given the class. Despite its independence assumption, the GNB classifier can be surprisingly robust and efficient, especially when the classes are well separated, as in the present case [59,60]. Finally, Log was chosen because it is another powerful model for binary classification that provides class membership probabilities. This model assumes a linear relationship between the predictor variables and the log-odds of the target class [61].

In this study, in order to find for each classifier and VI the best setting of hyperparameters (parameters that are set prior to the training process and cannot be directly learned from the data) capable of ensuring the best performance, the function GridSearchCV with scikit-learn module was used. For each combination of hyperparameters, the function trains the model using cross-validation and evaluates its performance. Adopting this technique across all VIs, each classifier employs the best possible combination of hyperparameters within a proposed grid, thereby significantly reducing the possibility that the final model performance is influenced by an arbitrary choice.

Support Vector Machine (SVM) Classifier

The SVM classifier is a versatile machine learning algorithm that is widely employed for both classification and regression tasks, and it operates by finding the hyperplane that best separates data points belonging to different classes in the feature space. The main strength of the SVM lies in its ability to handle linear and non-linear relationships in the data, making it a versatile choice for a wide range of applications, including image classification [84–86]. In this work, the *C-Support Vector Classification* algorithm based on *LIBSVM* [87] and integrated into Scikit-Learn Python module [88] was used. The combination of hyperparameters used by GridSearchCV function for this classifier was as follows: Kernel type `[rbf, linear]`; Parameter `C [1, 10, 100, 1000]`; Parameter `gamma [0.0003, 0.0004]`.

Gaussian Naive Bayes (GNB) Classifier

The GNB classifier is a probabilistic machine learning algorithm that is particularly well-suited for classification tasks where the goal is to assign an input data point to one of several predefined classes based on its features. It leverages conditional probability to make predictions by estimating the likelihood of a particular class given the observed features of an input. Despite its 'naive' assumption of independence among features,

which is often unrealistic in real-world scenarios, the GNB classifier has demonstrated effectiveness in various applications, especially when dealing with continuous and normally distributed data, and the Gaussian distribution simplifies the estimation of the probability density function. In context, the assumption is that the features within each class allow for efficient parameter estimation with limited training data. The classifier calculates the class posterior probabilities for a given input and assigns the data point to the class with the highest probability [89]. The hyperparameter tuning technique has not been applied to this classifier as it is non-parametric, but we modulated the portion of the largest variance of all features that is added to variances for calculation stability using a logarithmic (base = 10) space of 100 values, calculated from 0 to −9.

For GNB, we tested log-spaced priors for smoothing. Hence, within the selected algorithms, we performed systematic parameter optimization to ensure that each classifier uses the best set of hyperparameters.

Logistic (Log) Classifier

The logistic classifier is a widely used statistical method in the field of machine learning, and statistical modeling for binary classification problems is particularly well-suited for scenarios where the dependent variable is categorical and binary, meaning it has only two possible outcomes. The logistic classifier is employed to predict the probability that an instance belongs to a particular class.

Unlike linear regression, which predicts continuous outcomes, logistic regression models the probability of an event occurring.

The logistic model utilizes the logistic function [90], also known as the sigmoid function, to map any real-valued number into a range between 0 and 1. The combination of hyperparameters used by the GridSearchCV function for this classifier was as follows: Solver type `[newton-cg, lbfgs, liblinear]`; Parameter Penality `[none, l1, l2, elasticnet]`; Parameter `C [1, 10, 100, 1000]`; Parameter classweight `[balanced]`.

2.6. Training and Testing Data

Because the number of extracted pixels varied among tree crowns, it was necessary to standardize the array sizes containing the VI values. Specifically, for each crown and each VI, the arrays were sorted in descending order and truncated to 1710 samples (corresponding to pixels contained in the smallest crown) by retaining only the 1710 highest values. Then, the derived dataset, consisting of 67 tree crowns, was divided into two different subsets. Regarding the division criteria, a random state parameter = 10 (this parameter controls the shuffling applied to the data before applying the split, ensuring the reproducibility of the experiment) was adopted according to a stratified option. The first subset, the training dataset, was represented by 70% (44 samples out of 67) of the original dataset; the second subset, the remaining 30% (23 samples out of 67), made up the test set. Subsequently, the two datasets were further divided into features representing the independent variable (X train and X test) and features constituting the target values (the dependent variable; Y train and Y test). Furthermore, a stratified splitting method based on class labels was employed to ensure a balanced representation across both subsets. A label encoder transformation was also applied to the target values to normalize the original labels (asymptomatic and symptomatic) with values of 0 and 1. Specific procedures in Python language were developed for all data preprocessing operations, utilizing the scikit-learn library version 1.4.1 [88,91].

2.7. Model Evaluation

To evaluate the performances of each classifier applied to the eight different VIs, we derived several metrics calculated using the confusion matrices represented by the following four parameters:

- True positive (TP): represents the number of symptomatic trees correctly classified as symptomatic;
- True negative (TN): represents the number of asymptomatic trees correctly classified as asymptomatic;
- False positive (FP): represents the number of asymptomatic trees incorrectly classified as asymptomatic;
- False negative (FN): represents the number of symptomatic trees incorrectly classified as symptomatic.

The definition and formulas of the performance metrics are as follows:

Accuracy (ACC): represents the proportion of correct classified instances to the total number of classifications

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

Precision (P): represents the ratio of correctly predicted positive instances to the total predicted positive instances

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

Recall detection rate (R): represents the ratio of correctly predicted positive instances to the overall number of actual positive instances

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$F1_{Score}$ (F1s): represents the weighted average of precision and recall values and can be used as a single measure of performance of the test for the positive class

$$F1s = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (13)$$

In addition, the values of the Support (number of samples of the true response that lie in that symptomatic/asymptomatic class), Weighted Average (averaging the support-weighted mean per symptomatic/asymptomatic class), and Macro Average (averaging the unweighted mean per symptomatic/asymptomatic class) were calculated.

In this work, we applied a binary classification, and a 2 × 2 confusing matrix was used to describe the performance of machine learning classifiers (Table 5) [92].

Table 5. 2 × 2 confusing matrix used in this study to describe the performance of each classifier.

	Predictive Records	
Actual records	TP FN	FP TN

In order to facilitate replication, we have archived the relevant datasets (imagery, derived layers, annotated crown polygons, and model training) in a Zenodo repository and the data can be accessed upon request.

3. Results

Concerning the correlation between the VI mean value and LAI of each chestnut crown, highly significant correlations ($p < 0.001$) were found for all VIs except for MCARI (Figure 5). The correlations among all VIs and LAIs showed an R^2 higher than 0.53, while the maximum value (R^2 = 0.81) was recorded for ExGreenRed.

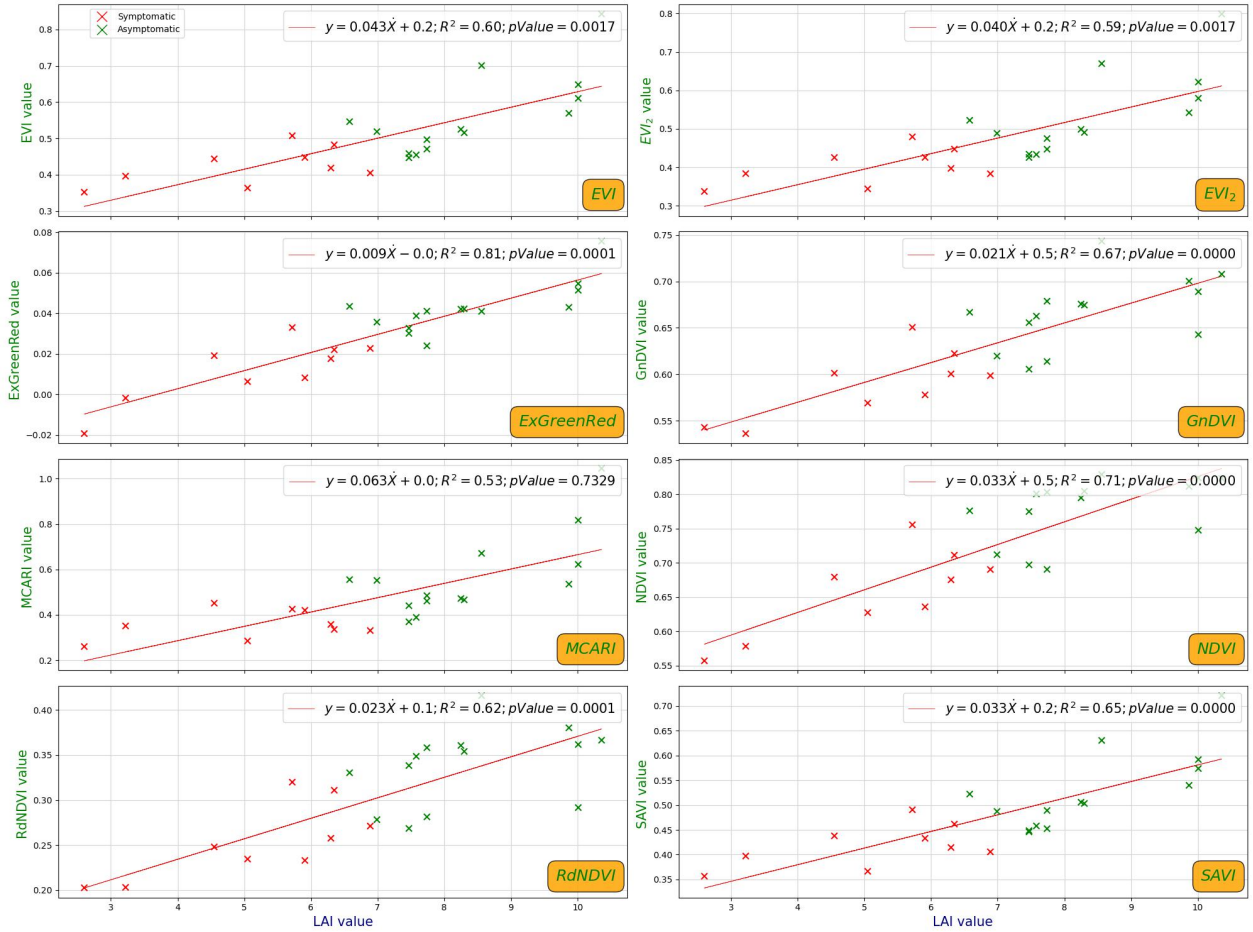
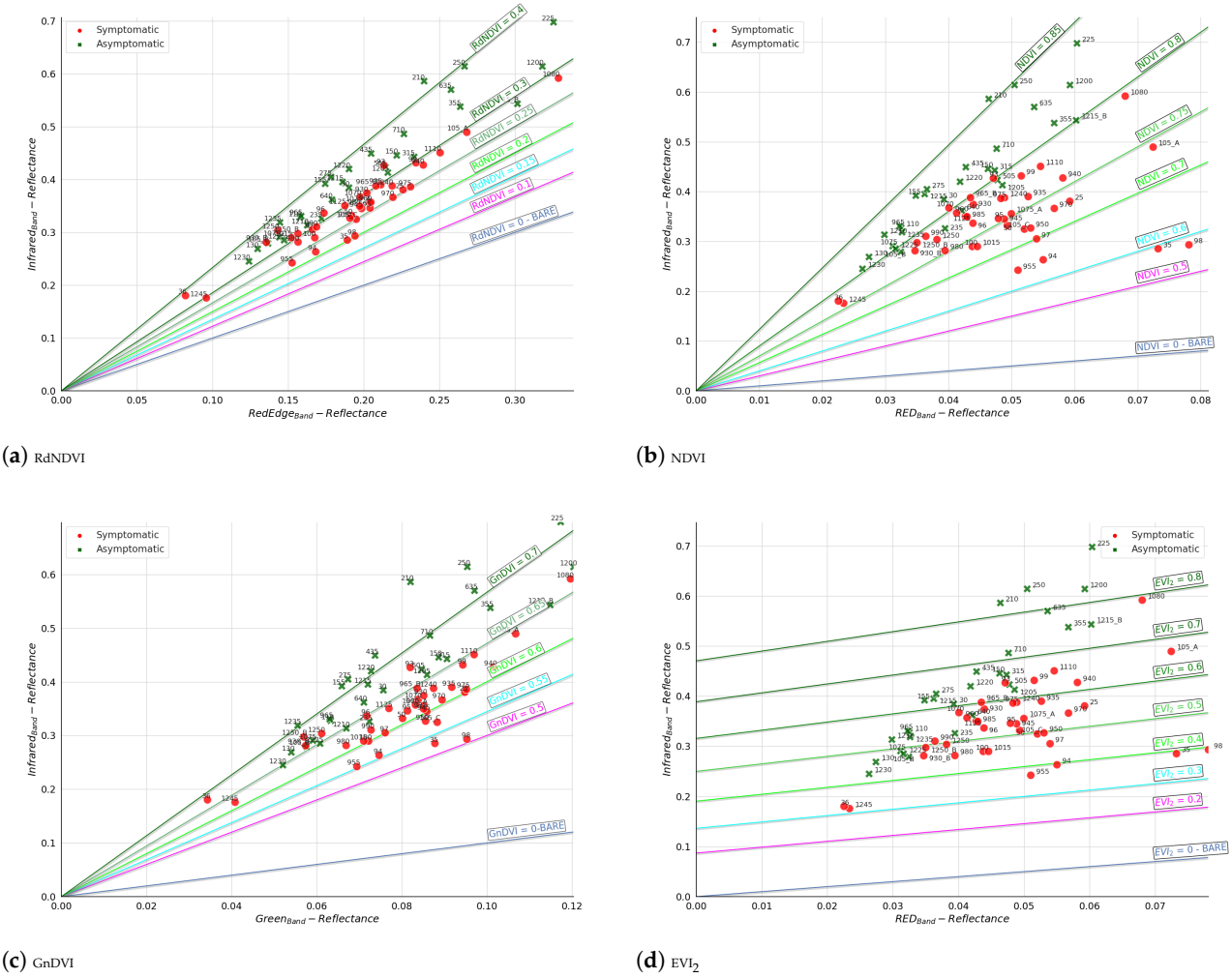


Figure 5. Linear regression between LAI and VI average values of chestnut crowns. Red = symptomatic trees; green = asymptomatic trees.

The isocline plots (Figure 6) highlighted that for three of four VIs derived by two spectral bands, it was possible to identify a correlation line to discriminate between symptomatic and asymptomatic trees (the R^2 were 0.8; 0.65–0.70; 0.3–0.4 for NDVI, GnDVI, and RdNDVI, respectively). The distribution of NDVI appeared sparser (values between 0.6 and 0.85) than those of RdNDVI and GnDVI (0.2–0.4 and 0.5–0.7, respectively) (Figure 6).

Instead, for EVI_2, it was not possible to identify a separation plane that discriminated between the health statuses of the plants.



(**a**) RdNDVI (**b**) NDVI

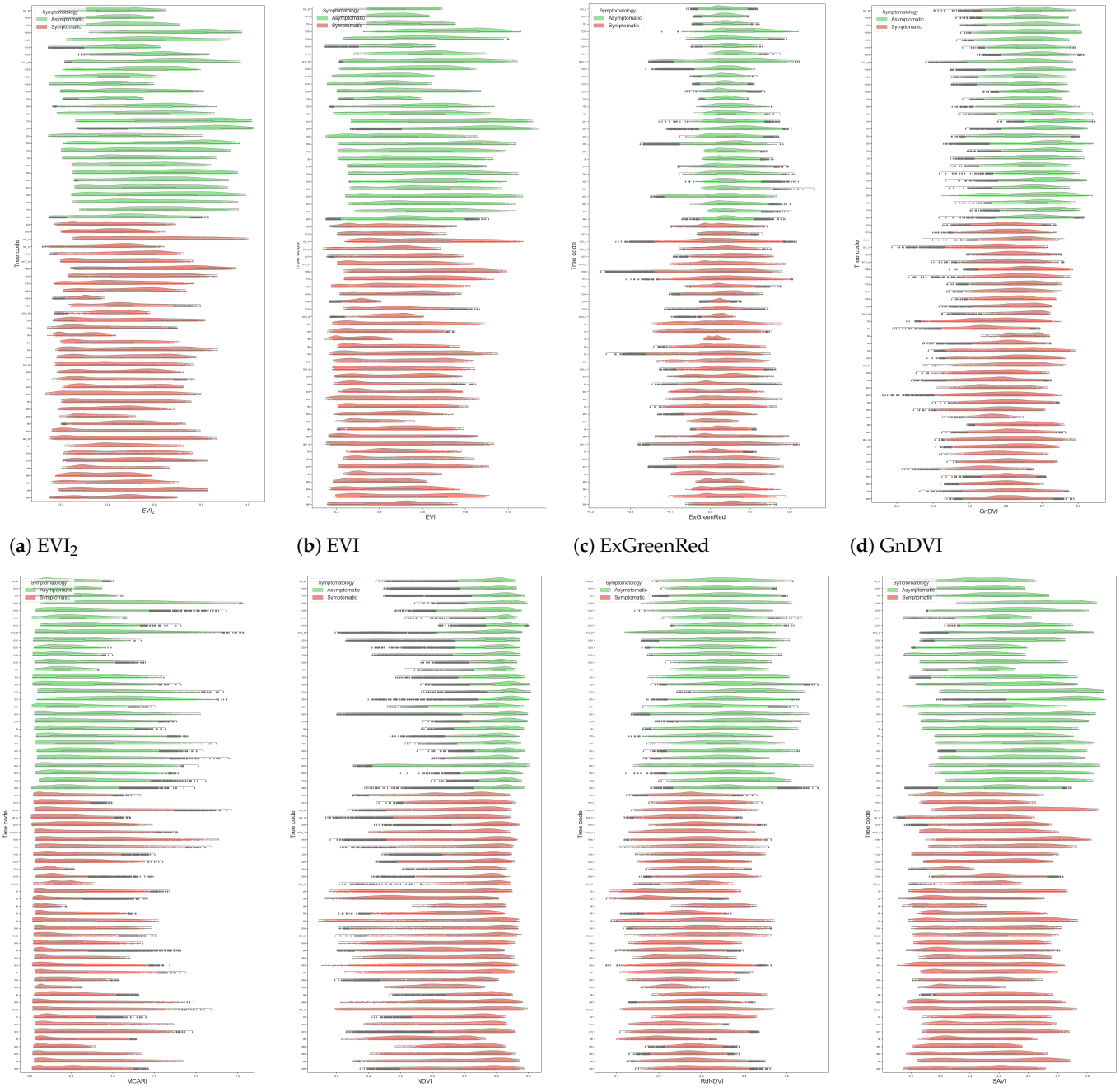
(**c**) GnDVI (**d**) EVI_2

Figure 6. Isocline plots of the four VIs computed by two spectral bands. Red = symptomatic trees; green = asymptomatic trees.

Considering each chestnut crown separately, for the pixel distribution for each VI in the cloud plots, asymptomatic trees showed a shift toward higher values, especially for EVI, NDVI, and RdNDVI (Figure 7b,f,g).

Similarly, the pixel distributions of VIs, shown by boxen plots and grouped for the health status of the trees, varied according to the tree health visual classification (Figure 8). In particular, NDVI showed a greater difference and a smaller interquartile range of asymptomatic trees compared to symptomatic ones (Figure 8f), while for MCARI, the boxen plot of both health statuses showed a similar shape (Figure 8e).

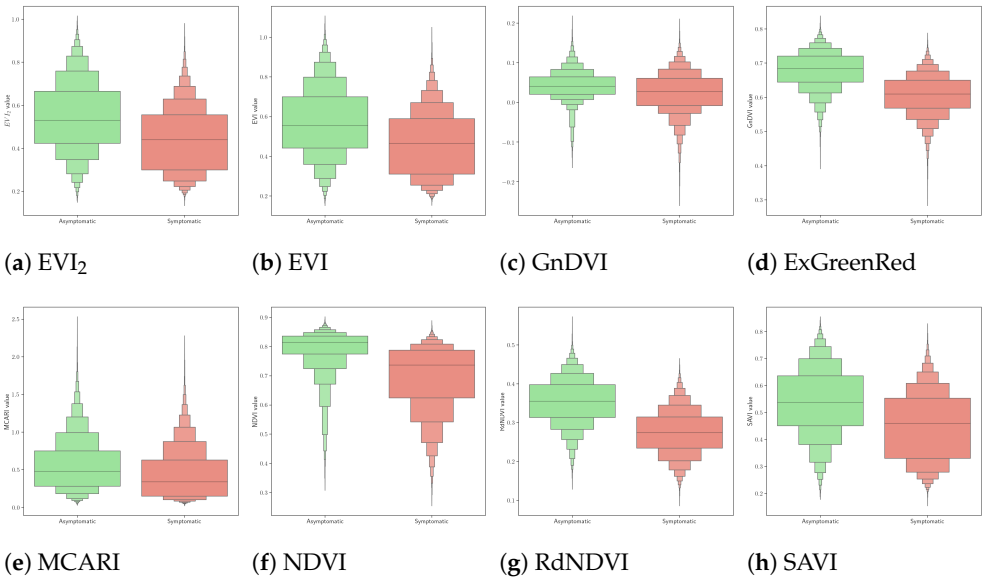
The PDF of the VIs grouped by health status showed a normal distribution except for SAVI, EVI, and EVI_2 for symptomatic trees, which were distributed in a bimodal pattern (Figure 9). NDVI exhibited the largest differences in terms of median value (0.74 for symptomatic trees and 0.81 for asymptomatic trees), while the ExGreenRed index exhibited the smallest ones (0.03 vs. 0.04).



(**a**) EVI_2 (**b**) EVI (**c**) ExGreenRed (**d**) GnDVI

(**e**) MCARI (**f**) NDVI (**g**) RdNDVI (**h**) SAVI

Figure 7. Cloud plots of the pixel distribution for each VI considering the crowns of chestnut trees separately. Red = symptomatic trees; green = asymptomatic trees.



(**a**) EVI_2 (**b**) EVI (**c**) GnDVI (**d**) ExGreenRed

(**e**) MCARI (**f**) NDVI (**g**) RdNDVI (**h**) SAVI

Figure 8. Boxen plots of the pixel of each VI value grouped by health status of the chestnut trees. Red = symptomatic trees; green = asymptomatic trees.

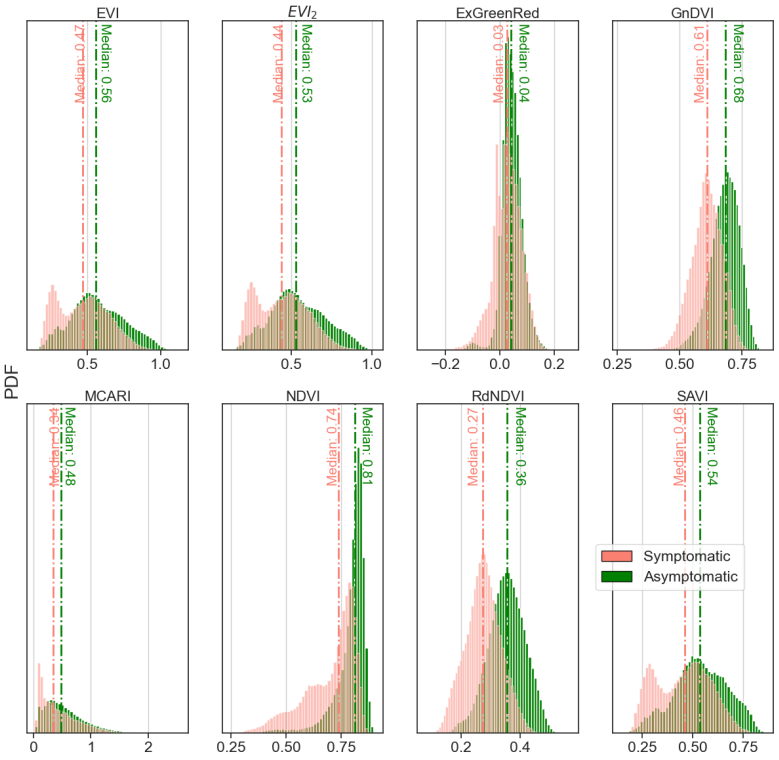
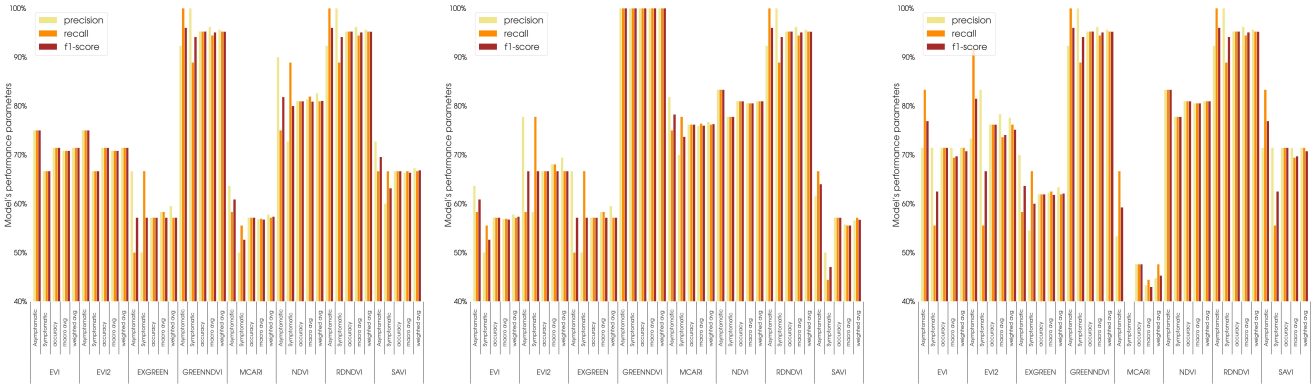


Figure 9. Probability density function (PDF) of each VI, grouped for chestnut health status. In the abscissas are the VI values; in ordinates, the relative frequencies [0–1]. Red = symptomatic trees; green = asymptomatic trees.

Classification Approach Performances

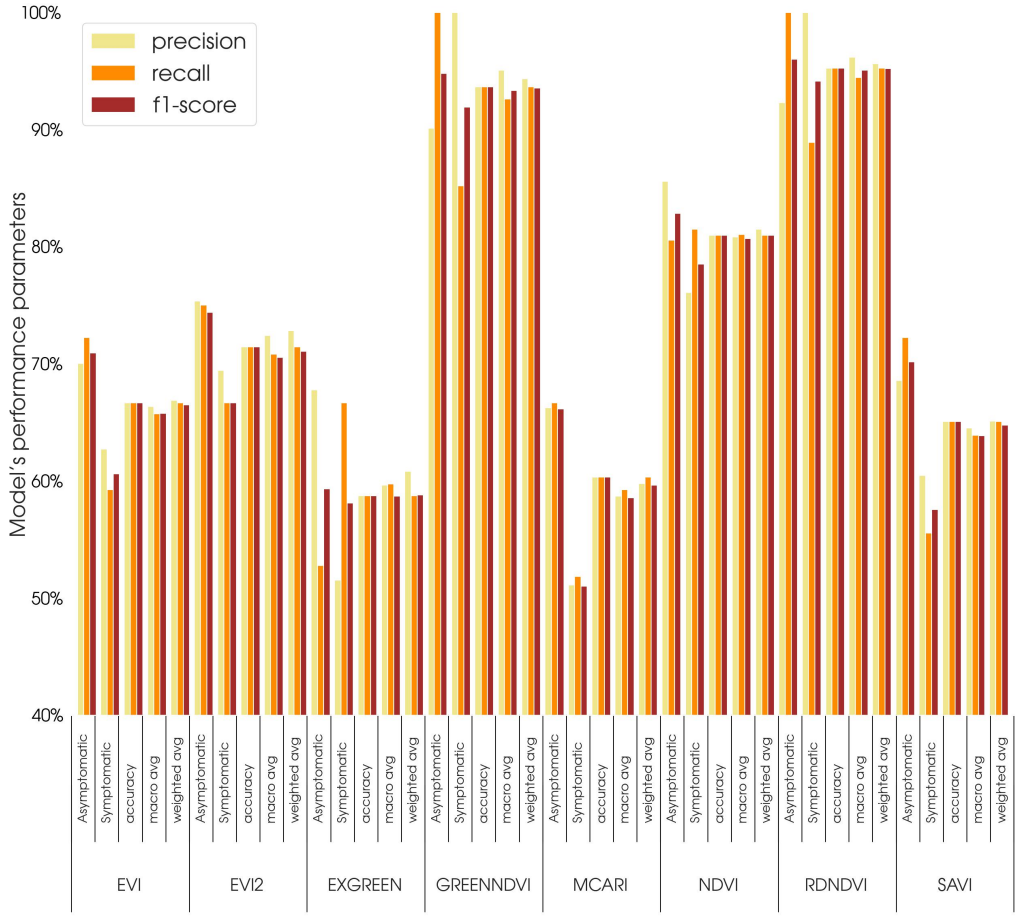
This section reports the results of the three classifiers applied to the eight VIs considered in this study. In Figure 10 and Table 6, the performances, expressed in terms of precision, recall, and F1-score, for the parameters of asymptomatic, symptomatic, accuracy, macro avg, and weighted avg are shown.



(**a**) Performance parameters related to GNB classifier.

(**b**) Performance parameters related to Log classifier.

(**c**) Performance parameters related to SVM classifier.



(**d**) Average performances by VIs

Figure 10. Performance parameters.

The outcome of the classifier's application to each VIs revealed that GnDVI and RdNDVI exhibited the best performance, obtaining precision values above 90%. For both GnDVI and RdNDVI, the best results of accuracy were obtained by applying the SVM and GNB classifiers (95.2%), while using the Log classifier, the accuracy decreased to 90.5% for GnDVI. Furthermore, for NDVI, which is the most widely used VI to monitor the state of vegetation health, the accuracy was valuable (above 80% for all classifiers), even if GNB seemed to better discriminate asymptomatic plants (precision of 90% compared to 72.7% related to symptomatic plants). The EVI, EVI_2, and SAVI performed at below 80% accuracy on average for all classifiers (66.7%, 71.5%, and 65.1%, respectively). The Log was less

accurate for these indices than the other two classifiers (71.4%, 76.2%, and 71.4% for SVM; 71.4%, 71.4%, and 66.7% for GNB; and 57.1%, 66.7%, and 57.1% for Log, respectively).

The MCARI index, on the other hand, showed the lowest performance with all classifiers: the accuracy was 47.6% for SVM, 76.2% for Log, and 57.1% for GNB. On average, the tendency of all classifiers was to classify trees as asymptomatic rather than symptomatic (average accuracy of 66.1% and 51.1%, respectively).

Table 6. Performance parameters for each of the eight VIs and for the three classifiers in terms of precision, recall, F1-score, and support.

		SVM				Log				GNB				Average			
VIs	**Parameters**	**Precision**	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
EVI	Asymptomatic	71.4%	83.3%	76.9%	12	63.6%	58.3%	60.9%	12	75.0%	75.0%	75.0%	12	70.0%	72.2%	70.9%	12
	Symptomatic	71.4%	55.6%	62.5%	9	50.0%	55.6%	52.6%	9	66.7%	66.7%	66.7%	9	62.7%	59.3%	60.6%	9
	Accuracy	71.4%	71.4%	71.4%	71.4%	57.1%	57.1%	57.1%	57.1%	71.4%	71.4%	71.4%	71.4%	66.7%	66.7%	66.7%	66.7%
	Macro avg	71.4%	69.4%	69.7%	21	56.8%	56.9%	56.8%	21	70.8%	70.8%	70.8%	21	66.4%	65.7%	65.8%	21
	Weighted avg	71.4%	71.4%	70.7%	21	57.8%	57.1%	57.3%	21	71.4%	71.4%	71.4%	21	66.9%	66.7%	66.5%	21
EVI_2	Asymptomatic	73.3%	91.7%	81.5%	12	77.8%	58.3%	66.7%	12	75.0%	75.0%	75.0%	12	75.4%	75.0%	74.4%	12
	Symptomatic	83.3%	55.6%	66.7%	9	58.3%	77.8%	66.7%	9	66.7%	66.7%	66.7%	9	69.4%	66.7%	66.7%	9
	Accuracy	76.2%	76.2%	76.2%	76.2%	66.7%	66.7%	66.7%	66.7%	71.4%	71.4%	71.4%	71.4%	71.4%	71.4%	71.4%	71.4%
	Macro avg	78.3%	73.6%	74.1%	21	68.1%	68.1%	66.7%	21	70.8%	70.8%	70.8%	21	72.4%	70.8%	70.5%	21
	Weighted avg	77.6%	76.2%	75.1%	21	69.4%	66.7%	66.7%	21	71.4%	71.4%	71.4%	21	72.8%	71.4%	71.1%	21
ExGreenRed	Asymptomatic	70.0%	58.3%	63.6%	12	66.7%	50.0%	57.1%	12	66.7%	50.0%	57.1%	12	67.8%	52.8%	59.3%	12
	Symptomatic	54.5%	66.7%	60.0%	9	50.0%	66.7%	57.1%	9	50.0%	66.7%	57.1%	9	51.5%	66.7%	58.1%	9
	Accuracy	61.9%	61.9%	61.9%	61.9%	57.1%	57.1%	57.1%	57.1%	57.1%	57.1%	57.1%	57.1%	58.7%	58.7%	58.7%	58.7%
	Macro avg	62.3%	62.5%	61.8%	21	58.3%	58.3%	57.1%	21	58.3%	58.3%	57.1%	21	59.6%	59.7%	58.7%	21
	Weighted avg	63.4%	61.9%	62.1%	21	59.5%	57.1%	57.1%	21	59.5%	57.1%	57.1%	21	60.8%	58.7%	58.8%	21
GnDVI	Asymptomatic	92.3%	100.0%	96.0%	12	85.7%	100.0%	92.3%	12	92.3%	100.0%	96.0%	12	90.1%	100.0%	94.8%	12
	Symptomatic	100.0%	88.9%	94.1%	9	100.0%	77.8%	87.5%	9	100.0%	88.9%	94.1%	9	100.0%	85.2%	91.9%	9
	Accuracy	95.2%	95.2%	95.2%	95.2%	90.5%	90.5%	90.5%	90.5%	95.2%	95.2%	95.2%	95.2%	93.7%	93.7%	93.7%	93.7%
	Macro avg	96.2%	94.4%	95.1%	21	92.9%	88.9%	89.9%	21	96.2%	94.4%	95.1%	21	95.1%	92.6%	93.3%	21
	Weighted avg	95.6%	95.2%	95.2%	21	91.8%	90.5%	90.2%	21	95.6%	95.2%	95.2%	21	94.3%	93.7%	93.5%	21
MCARI	Asymptomatic	53.3%	66.7%	59.3%	12	81.8%	75.0%	78.3%	12	63.6%	58.3%	60.9%	12	66.3%	66.7%	66.1%	12
	Symptomatic	33.3%	22.2%	26.7%	9	70.0%	77.8%	73.7%	9	50.0%	55.6%	52.6%	9	51.1%	51.9%	51.0%	9
	Accuracy	47.6%	47.6%	47.6%	47.6%	76.2%	76.2%	76.2%	76.2%	57.1%	57.1%	57.1%	57.1%	60.3%	60.3%	60.3%	60.3%
	Macro avg	43.3%	44.4%	43.0%	21	75.9%	76.4%	76.0%	21	56.8%	56.9%	56.8%	21	58.7%	59.3%	58.6%	21
	Weighted avg	44.8%	47.6%	45.3%	21	76.8%	76.2%	76.3%	21	57.8%	57.1%	57.3%	21	59.8%	60.3%	59.6%	21
NDVI	Asymptomatic	83.3%	83.3%	83.3%	12	83.3%	83.3%	83.3%	12	90.0%	75.0%	81.8%	12	85.6%	80.6%	82.8%	12
	Symptomatic	77.8%	77.8%	77.8%	9	77.8%	77.8%	77.8%	9	72.7%	88.9%	80.0%	9	76.1%	81.5%	78.5%	9
	Accuracy	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%	81.0%
	Macro avg	80.6%	80.6%	80.6%	21	80.6%	80.6%	80.6%	21	81.4%	81.9%	80.9%	21	80.8%	81.0%	80.7%	21
	Weighted avg	81.0%	81.0%	81.0%	21	81.0%	81.0%	81.0%	21	82.6%	81.0%	81.0%	21	81.5%	81.0%	81.0%	21
RdNDVI	Asymptomatic	92.3%	100.0%	96.0%	12	92.3%	100.0%	96.0%	12	92.3%	100.0%	96.0%	12	92.3%	100.0%	96.0%	12
	Symptomatic	100.0%	88.9%	94.1%	9	100.0%	88.9%	94.1%	9	100.0%	88.9%	94.1%	9	100.0%	88.9%	94.1%	9
	Accuracy	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%	95.2%
	Macro avg	96.2%	94.4%	95.1%	21	96.2%	94.4%	95.1%	21	96.2%	94.4%	95.1%	21	96.2%	94.4%	95.1%	21
	Weighted avg	95.6%	95.2%	95.2%	21	95.6%	95.2%	95.2%	21	95.6%	95.2%	95.2%	21	95.6%	95.2%	95.2%	21
SAVI	Asymptomatic	71.4%	83.3%	76.9%	12	61.5%	66.7%	64.0%	12	72.7%	66.7%	69.6%	12	68.6%	72.2%	70.2%	12
	Symptomatic	71.4%	55.6%	62.5%	9	50.0%	44.4%	47.1%	9	60.0%	66.7%	63.2%	9	60.5%	55.6%	57.6%	9
	Accuracy	71.4%	71.4%	71.4%	71.4%	57.1%	57.1%	57.1%	57.1%	66.7%	66.7%	66.7%	66.7%	65.1%	65.1%	65.1%	65.1%
	Macro avg	71.4%	69.4%	69.7%	21	55.8%	55.6%	55.5%	21	66.4%	66.7%	66.4%	21	64.5%	63.9%	63.9%	21
	Weighted avg	71.4%	71.4%	70.7%	21	56.6%	57.1%	56.7%	21	67.3%	66.7%	66.8%	21	65.1%	65.1%	64.8%	21

In Table 7, the overall results of the three classifiers are reported. Across all evaluated metrics, the classifiers demonstrated varying levels of performance. The SVM achieved the highest overall accuracy (75.0%) with precision and recall scores that favored the asymptomatic class (75.9% precision, 83.3% recall). In contrast, the GNB classifier attained an accuracy of 74.4%, with higher precision for asymptomatic individuals (78.5%) than symptomatic (70.8%), although its recall values were comparatively similar across health status classes (75.0% vs. 73.6%). Finally, the Log classifier showed the lowest overall accuracy (72.6%), with asymptomatic precision (76.6%) outpacing its symptomatic counterpart (69.5%), but showing similar recall (74.0% vs. 70.8%). Weighted average values further underscored the interplay between class distributions and performance, highlighting that all three classifiers exhibited comparable discriminative power in identifying asymptomatic and symptomatic trees.

Table 7. Overall accuracy parameters for the three models.

Classifier	Parameters	Precision	Recall	F1-Score	Support
GNB	Asymptomatic	78.5%	75.0%	76.4%	12
	Symptomatic	70.8%	73.6%	71.8%	9
	Accuracy	74.4%	74.4%	74.4%	74.4%
	Macro avg	74.6%	74.3%	74.1%	21
	Weighted avg	75.2%	74.4%	74.4%	21
Log	Asymptomatic	76.6%	74.0%	74.8%	12
	Symptomatic	69.5%	70.8%	69.6%	9
	Accuracy	72.6%	72.6%	72.6%	72.6%
	Macro avg	73.1%	72.4%	72.2%	21
	Weighted avg	73.6%	72.6%	72.6%	21
SVM	Asymptomatic	75.9%	83.3%	79.2%	12
	Symptomatic	74.0%	63.9%	68.0%	9
	Accuracy	75.0%	75.0%	75.0%	75.0%
	Macro avg	75.0%	73.6%	73.6%	21
	Weighted avg	75.1%	75.0%	74.4%	21

4. Discussion

By integrating the LAI with spectral indices, this study presents a holistic approach to chestnut tree health assessments, confirming the possibility of a more accurate detection of disease impacts and providing a foundation for precision forestry practices aimed at mitigating ink disease effects. The results highlight the promising application of multispectral UAV imagery and machine learning classifiers in assessing ink disease in chestnut orchards. The study confirmed that (a) chestnut trees' health status related to ink disease can be reliably described by LAI values (Figure 5), and (b) the accuracy values obtained confirm that all classification algorithms used in this work can recognize the health status of plants, in line with careful and accurate analysis conducted by experienced pathologists [93].

Indeed, a strong correlation emerged between the LAI values and the average VI values at the plant level, highlighting the LAI's potential as a complementary indicator for assessing chestnut tree health. Specifically, the correlation analysis between the LAI and various VIs showed statistically significant results for all indices, except for the MCARI. The highest correlation was observed with the ExGreenRed index, indicating its strong relationship with canopy coverage and vigor, though it performed less effectively in classification tasks, whereas NDVI and GnDVI demonstrated a robust dual utility in both VI–LAI correlation and tree health classification.

The LAI values further stratified tree health into symptomatic and asymptomatic categories. Symptomatic trees exhibited significantly reduced LAI values compared to asymptomatic trees, reflecting the impact of ink disease on canopy density and leaf health. The mean LAI for symptomatic trees was approximately 5.91, compared to 8.30 for asymptomatic trees. This marked difference highlights the utility of the LAI as an indicator of physiological stress in correlation with observable disease symptoms such as sparse foliage and reduced vegetation cover.

However, in this work, the correlation between the LAI and VIs was mainly used to validate the visual classification made by forest pathologists. The strong correlations suggested that LAI values can still give an indication of the physiological status of plants and could allow orchard managers to flag suspected trees for closer inspection or prophylactic treatment sooner, thus potentially limiting the spread of *Phytophthora* pathogens. The relationship between the LAI and VIs was further supported by the PDF plots (Figure 9), which revealed distinct patterns between symptomatic and asymptomatic groups of trees. Asymptomatic trees consistently exhibited higher VI values with narrower interquartile ranges in the corresponding VI distributions, especially for NDVI and GrNDVI. In contrast, symptomatic trees displayed broader distributions and, in some cases, bimodal patterns,

such as those seen with SAVI and EVI_2. These findings suggest greater variability in the health responses of symptomatic trees, potentially due to the varying severity of ink disease among affected individuals.

These results agree with the findings of [94], where the integration of vegetation indices and geometric parameters in chestnut groves yielded an 85% R^2 value for LAI predictions, demonstrating the effectiveness of combining multispectral data with LAI.

Furthermore, the isocline plots for indices derived from two spectral bands, especially for NDVI, RdNDVI, and GnDVI, showed well-defined linear trends effectively separating symptomatic from asymptomatic trees. The ability to establish such separation underscores the complementary role of the LAI in enhancing the interpretability and precision of spectral data in health classification. This alignment of LAI and spectral indices reinforces their combined potential to diagnose and monitor disease progression in forest ecosystems. However, it is important to note that certain indices, like SAVI, despite their strong LAI correlations (Figure 5), underperformed in health classification tasks. This discrepancy may be attributed to these indices' reduced sensitivity to specific stress-related spectral signatures or structural variations in diseased canopies, emphasizing that not all high-LAI-correlating indices are equally effective for disease classification. In summary, LAI has proven itself to be a critical metric for validating and enhancing spectral indices in the context of disease monitoring. Its correlation with indices such as NDVI, GnDVI, and RdNDVI confirms its role as a robust physiological indicator, aligning with the observed differences in canopy density and health status.

Among the eight VIs analyzed, NDVI and GnDVI stood out for their ability to effectively discriminate between symptomatic and asymptomatic chestnut trees. These indices achieved the highest classification accuracy, with GnDVI showing higher performance when paired with SVM and GNB classifiers. Similarly, NDVI demonstrated robust results, maintaining an accuracy of 81% across all classifiers. The Log classifier, while effective for NDVI and GnDVI, displayed slightly reduced performance for NDVI. This difference underscores the sensitivity of classifiers to the spectral properties of vegetation indices. Other indices, such as MCARI and EVI, exhibited significantly lower classification accuracy.

For MCARI, the observed decline in accuracy appears partly attributable to the index's core focus on chlorophyll absorption in the red spectral region. However, MCARI also incorporates reflectance in the green channel, which can be shaped by canopy architecture and environmental factors unrelated to incipient disease symptoms. In certain cases, large or overlapping canopies may attenuate the specific red–green contrast upon which MCARI relies, thereby diluting its capacity to detect signs of stress [45]. Consequently, structural or external elements may overshadow changes in chlorophyll and reduce MCARI's overall sensitivity to disease identification [95].

With regard to EVI, the lower accuracy is likely tied to its additional coefficients designed to correct for aerosol scattering and soil reflectance. While EVI often performs well in highly dense canopies (e.g., tropical forests), its sensitivity to subtle stress signals may diminish if the vegetation is not extremely dense or if the pathogen primarily induces localized leaf loss or discoloration. In such circumstances, the index's corrective parameters may inadvertently mask the fine-scale spectral variations needed for disease detection [96].

This study aligns with broader applications of UAV-based monitoring in forestry, as discussed in systematic reviews highlighting the advantages of UAVs in capturing high-resolution data for pest and disease detection [28]. The demonstrated ability to differentiate symptomatic from asymptomatic chestnut trees through spectral signatures is an advancement in detecting tree-level attributes as reliable proxies for tree vigor and stress [53].

The RdNDVI also performed exceptionally well, achieving a high accuracy across all classifiers. RdNDVI's utility likely stems from its sensitivity to physiological changes in vegetation linked to stress. This index, along with GnDVI and NDVI, exhibited greater consistency and lower interquartile ranges in pixel value distributions for asymptomatic trees, reflecting its reliability in detecting health differences. In contrast, indices such as EVI_2 and SAVI struggled to separate symptomatic from asymptomatic crowns effectively, with accuracy values consistently low and bimodal patterns observed in their PDF.

These results suggest that MCARI, while valuable for other applications, may lack the robustness needed for precise health discrimination in chestnut trees affected by ink disease. Similarly, EVI showed a reduced ability to capture the subtle spectral variations between symptomatic and asymptomatic trees. The limitations of those underperforming VIs are consistent with studies that highlight variability in the effectiveness of vegetation indices depending on canopy structure and disease severity [94,97].

An interesting general observation of the classifier performance metrics was their ability to distinguish asymptomatic trees more accurately than symptomatic ones (higher precision and higher F1 score). This trend suggested that asymptomatic trees may exhibit more stable and consistent spectral characteristics, making them easier to classify accurately.

A further interesting aspect to be highlighted is the bimodal pattern of SAVI, EVI, and EVI_2 of symptomatic plants that could be explained by the fact these VIs are derived, considering a correction factor of the soil component that reverberates on the PDF.

Furthermore, the distribution of VI pixel values illustrated key differences in their ability to represent tree health. NDVI and RdNDVI exhibited significant shifts in their median values between symptomatic and asymptomatic trees. Conversely, indices like ExGreenRed showed minimal differences in median values among the two trees' health status, highlighting their limited discriminative capacity. The bimodal distributions observed in SAVI and EVI_2 for symptomatic trees further reflect the challenges in their application for trees' health classification, likely due to greater variability in the stress response of the crowns of infected trees.

The support vector machine and Gaussian Naive Bayes classifiers generally outperform the logistic model across most of the vegetation indices (VIs) (Table 6). Notably, both SVM and GNB achieve the highest overall accuracy (95.2%) with GnDVI and RdNDVI, indicating their superior ability to distinguish symptomatic from asymptomatic chestnut trees. These two indices, which exploit the near-infrared band, appear especially sensitive to subtle canopy changes brought on by ink disease, allowing the SVM and GNB to attain both high precision (low false positives) and recall (low false negatives). While logistic regression also provides competitive results for GnDVI and RdNDVI, it tends to lag slightly behind in terms of accuracy and F1-score, suggesting it is more sensitive to variation in the data. In contrast, for indices like MCARI, EVI, SAVI, and EVI_2, performance dropped more noticeably in the Log model. Overall, SVM and GNB demonstrate stronger consistency, particularly with the most informative indices, which makes them better suited for the detection of ink disease in this setting. Meanwhile, the Log model may still be viable for simpler or more linearly separable data but it shows weaknesses when the spectral signatures overlap or when disease symptoms exhibit greater variability. Therefore, considering accuracy, precision, recall, and F1-score together, SVM and GNB emerge as the more reliable choices for the classification of symptomatic versus asymptomatic chestnut trees in the orchard.

All three classifiers—GNB, Log, and SVM—exhibited comparable and, in some cases, nearly identical classification accuracies, particularly when applied to vegetation indices (VIs) such as GnDVI, RdNDVI, and NDVI. This convergence is largely attributable to the nature of the dataset and the inherent structure of the feature space:

- The VIs used (e.g., GnDVI and RdNDVI) demonstrated strong discriminatory power between symptomatic and asymptomatic classes. As shown in our isocline and distribution plots (Figure 6), these VIs produced relatively linearly separable or well-clustered data distributions;
- GNB assumes feature independence and Gaussian-distributed data, which is not strictly true for most real-world datasets. However, in our case, the VI values—especially at crown level—approximated unimodal and symmetric distributions, fulfilling GNB's assumptions reasonably well;
- The logistic classifier models the linear decision boundary in log-odds space, which works effectively when features correlate linearly with class probability. Our VIs, particularly those built on normalized differences (e.g., NDVI), exhibited such relationships, as corroborated by strong correlations with LAI and clear VI value ranges across health classes;
- The SVM, especially with linear or RBF kernels, is robust in both linearly and non-linearly separable contexts, but when the data are already separable (as in our case), its decision surface closely aligns with that of logistic regression and even GNB.

Overall, this study underscores the importance of selecting appropriate VIs and classifiers tailored to the physiological and spectral nuances of the target vegetation. GnDVI and RdNDVI, when paired with advanced machine learning models like SVM and GNB, emerged as the most reliable tools for detecting ink disease in chestnut trees. By contrast, the performance of MCARI and SAVI highlights the need for cautious interpretation and potential refinement when applying less robust indices in similar contexts. These findings provide valuable insights into optimizing remote sensing and machine learning methodologies for precision agriculture and forest health monitoring, offering actionable pathways for disease detection and resource management in affected ecosystems.

While the present study demonstrated robust results and high classification accuracy in differentiating symptomatic from asymptomatic chestnut trees affected by ink disease, we acknowledge the following limitations: sample size imbalance, our dataset consisted of an unequal number of symptomatic (38 trees) and asymptomatic (29 trees) samples, potentially introducing bias towards the majority class during model training and evaluation phases; limited number of LAI measurements: LAI measurements were limited to 10 trees per health status category, raising concerns about representativeness and statistical robustness in fully capturing variability across the orchard; single location and timeframe: data collection was limited to one orchard location and a single growing season, which may constrain the generalizability of the findings across different environmental conditions, seasons, and geographic locations. To comprehensively address and mitigate the aforementioned limitations, we propose several methodological enhancements in future research: increase field data collection, increase the number of LAI measurements per group, and incorporate data from multiple locations and across different timeframes, seasons, and conditions to enhance data representativeness and generalizability.

5. Conclusions

This study provides a significant step forward in integrating advanced remote sensing techniques and machine learning approaches for the detection and management of ink disease in chestnut orchards. Based on our results, the following key findings were identified:

- The spectral indices GnDVI and RdNDVI demonstrated the highest effectiveness, more than NDVI, and achieved classification accuracies up to 95.2% when combined with the SVM and GNB classifiers. These indices effectively captured physiological changes associated with ink disease symptoms and underscore the potential of high-resolution UAV imagery in achieving accurate tree health assessments;

- Significant correlations ($p < 0.001$) were observed between LAI and most vegetation indices, confirming LAI's value as a reliable physiological proxy for validating spectral assessments of chestnut tree health;
- Limitations of certain vegetation indices: indices such as MCARI and SAVI showed comparatively limited discriminatory power, highlighting the need for the careful selection of vegetation indices that are specifically tailored to reflect subtle physiological changes due to disease.

These findings are particularly valuable for large-scale forestry management, where the detection of stressors like ink disease is critical for implementing timely interventions. The study's innovative approach of combining vegetation indices with LAI measurements further enhances the reliability of these tools, offering a dual framework for spectral and structural analysis. The clear differentiation between symptomatic and asymptomatic trees achieved in this research establishes a solid foundation for scalable applications in forest health monitoring. However, the variability observed in the performance of certain indices, such as MCARI and ExGreenRed highlights the need for further refinement of vegetation indices tailored to specific stress responses. This may be because the MCARI is an expression of chlorophyll absorption, and there are variations in chlorophyll concentrations; therefore, in adult plants, the leaf cover expressed by the LAI is wide but not closely correlated with an equally efficient photosynthetic activity.

Future studies could focus on exploring additional indices or integrating hyperspectral data, which may capture subtle physiological variations more effectively. Moreover, expanding this approach to other tree species and environmental conditions would provide broader applicability and robustness. Advancements in machine learning, such as deep learning models, could also be leveraged to improve classification accuracy and adaptability to complex datasets. The prospective integration of these techniques into real-time monitoring systems would represent a promising avenue for sustainable forest management. By automating disease detection and incorporating predictive models, forest managers could transition from reactive to proactive strategies, minimizing disease spread and associated economic losses. Furthermore, linking these methods with ecological data on climate, soil, and biodiversity could provide a more comprehensive understanding of disease dynamics, fostering resilience in forestry ecosystems. This study sets the stage for a new era of precision forestry, where technology and data-driven insights converge to address pressing challenges in forest health and sustainability.

Author Contributions: Conceptualization, L.A. and G.D.R.; methodology, L.A., R.D. and G.D.R.; validation, L.A.; formal analysis, L.A. and M.C.; investigation, L.A., R.D., G.E., A.M., A.F., L.B. and G.D.R.; data field collection, L.A., A.F., G.E., D.P., A.M., L.B., S.B., N.S. and G.D.R.; data curation, L.A., D.P., M.C., N.S. and G.D.R.; writing—original draft preparation, L.A., G.E. and G.D.R.; writing—review and editing, L.A., R.D., G.E., A.M., S.B., A.F., M.C. and G.D.R.; supervision, L.A. and G.D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by LIFE MycoRestore project grant number LIFE 18/CCA/ES/001110.

Data Availability Statement: Data is contained within the article.

Acknowledgments: This paper was prepared in the frame of the MycoRestore project, funded by the European Union's LIFE program under grant agreement No. LIFE/18/CCA/001110. We are grateful to Giuseppe Salieri, "Azienda Agricola Le Casine" in San Godenzo, for hosting the demonstrative project area.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

ACC	accuracy
AGL	above-ground levels
B_{475}	$blue_{475nm\ wavelength}$
C3S	Copernicus Climate Change Service
DSM	digital surface model
DTM	digital terrain model
ECMWF	European Center for Medium-Range Weather Forecasts
EPSG	European Petroleum Survey Group
EU	European Union
EVI	enhanced vegetation index
EVI_2	enhanced vegetation index
ExGreenRed	excess green-excess red index
F1s	$F1_{Score}$
FN	false negative
FP	false positive
G_{560}	$green_{560nm\ wavelength}$
GNB	Gaussian Naive Bayes
GnDVI	green normalized difference vegetation
GNSS	Global Navigation Satellite System
GSD	ground sample distance
INFC	National Inventory of Forests and forest Carbon Pools
LAI	leaf area index
Log	logistic
MCARI	modified chlorophyll absorption in reflectance index
ML	machine learning
NDVI	normalized difference vegetation index
NiR_{840}	$near\ infrared_{840nm\ wavelength}$
P	precision
PDF	probability density function
R	recall detection rate
R_{668}	$red_{668nm\ wavelength}$
RdNDVI	red-edge normalized difference vegetation
RE_{717}	$red\ edge_{717nm\ wavelength}$
RTK	real-time kinematic positioning
SAR	synthetic aperture radar
SAVI	enhanced vegetation index
SfM	structure from motion
SVM	support vector machines
TN	true negative
TP	true positive
UAV	unmanned aerial vehicle
UID	unique identifier number
VIs	vegetation indices
VTOL	vertical take-off and landing

References

1. Vettraino, A.M.; Morel, O.; Perlerou, C.; Robin, C.; Diamandis, S.; Vannini, A. Occurrence and distribution of Phytophthora species in European chestnut stands, and their association with Ink Disease and crown decline. *Eur. J. Plant Pathol.* **2005**, *111*, 169–180. [CrossRef]
2. Turchetti, T.; Maresi, G. *Biological Control and Management of Chestnut Diseases*; Springer: Dordrecht, The Nertherlands, 2008; pp. 85–118. [CrossRef]

3. Conedera, M.; Krebs, P.; Gehring, E.; Wunder, J.; Hülsmann, L.; Abegg, M.; Maringer, J. How future-proof is Sweet chestnut (*Castanea sativa*) in a global change context? *For. Ecol. Manag.* **2021**, *494*, 119320. [CrossRef]
4. Di Renzo, L.; Bianchi, A.; De Lorenzo, A. Chestnut Fruits: Nutritional Value and New Products. In Proceedings of the Chestnut (*Castanea sativa*): A Multipurpose European Tree, Bruxelles, Belgium, 30 September 2010.
5. Bruzzese, S.; Blanc, S.; Brun, F. The Chestnut tree: A resource for the socio-economic revival of inland areas in a bio-economy perspective. In Proceedings of the Social and Ecological Value Added of Small-Scale Forestry to the Bio-Economy, Bolzano, Italy, 7–8 October 2020; p. 1. [CrossRef]
6. Vannini, A.; Vettraino, A.M. Ink disease in chestnuts: Impact on the European chestnut. *For. Snow Landsc. Res.* **2001**, *76*, 345–350.
7. Vettraino, A.M.; Natili, G.; Anselmi, N.; Vannini, A. Recovery and pathogenicity of Phytophthora species associated with a resurgence of ink disease in *Castanea sativa* Italy. *Plant Pathol.* **2001**, *50*, 90–96. [CrossRef]
8. Jung, T.; Durán, A.; von Stowasser, E.S.; Schena, L.; Mosca, S.; Fajardo, S.; González, M.; Ortega, A.D.N.; Bakonyi, J.; Seress, D.; et al. Diversity of Phytophthora species in Valdivian rainforests and association with severe dieback symptoms. *For. Pathol.* **2018**, *48*, e12443. [CrossRef]
9. Robin, C.; Morel, O.; Vettraino, A.M.; Perlerou, C.; Diamandis, S.; Vannini, A. Genetic variation in susceptibility to Phytophthora cambivora in European chestnut (*Castanea sativa*). *For. Ecol. Manag.* **2006**, *226*, 199–207. [CrossRef]
10. Şimşek, S.A.; Katrcoğlu, Y.Z.; Maden, S. Phytophthora Root Rot Diseases Occurring on Forest, Parks, and Nurseries in Turkey and Their Control Measures. *Turk. J. Agric.-Food Sci. Technol.* **2018**, *6*, 770–782. [CrossRef]
11. Frascella, A.; Sarrocco, S.; Mello, A.; Venice, F.; Salvatici, C.; Danti, R.; Emiliani, G.; Barberini, S.; Rocca, G.D. Biocontrol of Phytophthora xcambivora on *Castanea sativa*: Selection of Local *Trichoderma* spp. Isolates for the Management of Ink Disease. *Forests* **2022**, *13*, 1065. [CrossRef]
12. Carratore, R.D.; Aghayeva, D.N.; Ali-zade, V.M.; Bartolini, P.; Rocca, G.D.; Emiliani, G.; Pepori, A.; Podda, A.; Maserti, B. Detection of Cryphonectria hypovirus 1 in Cryphonectria parasitica isolates from Azerbaijan. *For. Pathol.* **2021**, *51*, e12718. [CrossRef]
13. Prospero, S.; Rigling, D. Invasion genetics of the chestnut blight fungus Cryphonectria parasitica in Switzerland. *Phytopathology* **2012**, *102*, 73–82. [CrossRef]
14. Venice, F.; Vizzini, A.; Frascella, A.; Emiliani, G.; Danti, R.; Rocca, G.D.; Mello, A. Localized reshaping of the fungal community in response to a forest fungal pathogen reveals resilience of Mediterranean mycobiota. *Sci. Total Environ.* **2021**, *800*, 149582. [CrossRef]
15. Erwin, D.C.; Ribeiro, O.K. *Phytophthora Diseases Worldwide*; American Phytopathological Society (APS Press): St. Paul, MN, USA, 1996.
16. Vannini, A.; Vettraino, A. Phytophthora cambivora. *Forest Phytophthoras* **2011**, *1*. [CrossRef]
17. Blom, J.M.; Vannini, A.; Vettraino, A.M.; Hale, M.D.; Godbold, D.L. Ectomycorrhizal community structure in a healthy and a Phytophthora-infected chestnut (*Castanea sativa* Mill.) stand in central Italy. *Mycorrhiza* **2009**, *20*, 25–38. [CrossRef] [PubMed]
18. Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Comput. Electron. Agric.* **2003**, *37*, 127–139. [CrossRef]
19. Prospero, S.; Heinz, M.; Augustiny, E.; Chen, Y.; Engelbrecht, J.; Fonti, M.; Hoste, A.; Ruffner, B.; Sigrist, R.; Van Den Berg, N.; et al. Distribution, causal agents, and infection dynamic of emerging ink disease of sweet chestnut in Southern Switzerland. *Environ. Microbiol.* **2023**, *25*, 2250–2265. [CrossRef]
20. Marzocchi, G.; Maresi, G.; Luchi, N.; Pecori, F.; Gionni, A.; Longa, C.M.O.; Pezzi, G.; Ferretti, F. 85 years counteracting an invasion: Chestnut ecosystems and landscapes survival against ink disease. *Biol. Invasions* **2024**, *26*, 2049–2062. [CrossRef]
21. Șandric, I.; Irimia, R.; Petropoulos, G.P.; Anand, A.; Srivastava, P.K.; Pleșoianu, A.; Faraslis, I.; Stateras, D.; Kalivas, D. Tree's detection & health's assessment from ultra-high resolution UAV imagery and deep learning. *Geocarto Int.* **2022**, *37*, 10459–10479. [CrossRef]
22. Iost Filho, F.H.; Heldens, W.B.; Kong, Z.; De Lange, E.S. Drones: Innovative Technology for Use in Precision Pest Management. *J. Econ. Entomol.* **2020**, *113*, 1–25. [CrossRef]
23. Ecke, S.; Stehr, F.; Frey, J.; Tiede, D.; Dempewolf, J.; Klemmt, H.J.; Endres, E.; Seifert, T. Towards operational UAV-based forest health monitoring: Species identification and crown condition assessment by means of deep learning. *Comput. Electron. Agric.* **2024**, *219*, 108785. [CrossRef]
24. De Luca, G.; Modica, G.; Silva, J.M.N.; Praticò, S.; Pereira, J.M. Assessing tree crown fire damage integrating linear spectral mixture analysis and supervised machine learning on Sentinel-2 imagery. *Int. J. Digit. Earth* **2023**, *16*, 3162–3198. [CrossRef]
25. Widjaja Putra, B.T.; Soni, P. Evaluating NIR-Red and NIR-Red edge external filters with digital cameras for assessing vegetation indices under different illumination. *Infrared Phys. Technol.* **2017**, *81*, 148–156. [CrossRef]
26. Fraser, B.T.; Congalton, R.G. Monitoring Fine-Scale Forest Health Using Unmanned Aerial Systems (UAS) Multispectral Models. *Remote Sens.* **2021**, *13*, 4873. [CrossRef]

27. Ollinger, S.V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* **2011**, *189*, 375–394. [CrossRef] [PubMed]
28. Duarte, A.; Borralho, N.; Cabral, P.; Caetano, M. Recent advances in forest insect pests and diseases monitoring using UAV-based data: A systematic review. *Forests* **2022**, *13*, 911. [CrossRef]
29. Cotrozzi, L. Spectroscopic detection of forest diseases: A review (1970–2020). *J. For. Res.* **2022**, *33*, 21–38. [CrossRef]
30. Stone, C.; Mohammed, C. Application of remote sensing technologies for assessing planted forests damaged by insect pests and fungal pathogens: A review. *Curr. For. Rep.* **2017**, *3*, 75–92. [CrossRef]
31. Yin, D.; Cai, Y.; Li, Y.; Yuan, W.; Zhao, Z. Assessment of the Health Status of Old Trees of Platycladus orientalis L. Using UAV Multispectral Imagery. *Drones* **2024**, *8*, 91. [CrossRef]
32. Abdulridha, J.; Ehsani, R.; Castro, A.D. Detection and differentiation between laurel wilt disease, phytophthora disease, and salinity damage using a hyperspectral sensing technique. *Agriculture* **2016**, *6*, 56. [CrossRef]
33. Sandino, J.; Pegg, G.; Gonzalez, F.; Smith, G. Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors* **2018**, *18*, 944. [CrossRef]
34. Gold, K.M.; Townsend, P.A.; Chlus, A.; Herrmann, I.; Couture, J.J.; Larson, E.R.; Gevens, A.J. Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sens.* **2020**, *12*, 286. [CrossRef]
35. Appeltans, S.; Pieters, J.G.; Mouazen, A.M. Potential of laboratory hyperspectral data for in-field detection of Phytophthora infestans on potato. *Precis. Agric.* **2022**, *23*, 876–893. [CrossRef]
36. Wavrek, M.T.; Carr, E.; Jean-Philippe, S.; McKinney, M.L. Drone remote sensing in urban forest management: A case study. *Urban For. Urban Green.* **2023**, *86*, 127978. [CrossRef]
37. Newby, Z.; Murphy, R.J.; Guest, D.I.; Ramp, D.; Liew, E.C.Y. Detecting symptoms of Phytophthora cinnamomi infection in Australian native vegetation using reflectance spectrometry: Complex effects of water stress and species susceptibility. *Australas. Plant Pathol.* **2019**, *48*, 409–424. [CrossRef]
38. Hornero, A.; Zarco-Tejada, P.J.; Quero, J.L.; North, P.R.J.; Ruiz-Gómez, F.J.; Sánchez-Cuesta, R.; Hernández-Clemente, R. Modelling hyperspectral-and thermal-based plant traits for the early detection of Phytophthora-induced symptoms in oak decline. *Remote Sens. Environ.* **2021**, *263*, 112570. [CrossRef]
39. Croeser, L.; Admiraal, R.; Barber, P.; Burgess, T.I.; Hardy, G.E.S.J. Reflectance spectroscopy to characterize the response of Corymbia calophylla to Phytophthora root rot and waterlogging stress. *Forestry* **2022**, *95*, 312–330. [CrossRef]
40. Gomes-Laranjo, J.; Araújo-Alves, J.; Ferreira-Cardoso, J.; Pimentel-Pereira, M.; Abreu, C.; Torres-Pereira, J. Effect of chestnut ink disease on photosynthetic performance. *J. Phytopathol.* **2004**, *152*, 138–144. [CrossRef]
41. Dinis, L.T.; Peixoto, F.; Zhang, C.; Martins, L.; Costa, R.; Gomes-Laranjo, J. Physiological and biochemical changes in resistant and sensitive chestnut (Castanea) plantlets after inoculation with Phytophthora cinnamomi. *Physiol. Mol. Plant Pathol.* **2011**, *75*, 146–156. [CrossRef]
42. Li, S.; Xu, L.; Jing, Y.; Yin, H.; Li, X.; Guan, X. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102640. [CrossRef]
43. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
44. Gao, S.; Yan, K.; Liu, J.; Pu, J.; Zou, D.; Qi, J.; Mu, X.; Yan, G. Assessment of remote-sensed vegetation indices for estimating forest chlorophyll concentration. *Ecol. Indic.* **2024**, *162*, 112001. [CrossRef]
45. Daughtry, C. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [CrossRef]
46. Huete, A. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
47. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.; Gao, X.; Ferreira, L. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
48. Jiang, Z.; Huete, A.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [CrossRef]
49. Gill, K.S.; Anand, V.; Chauhan, R.; Verma, G.; Gupta, R. Agricultural Pests Classification Using Deep Convolutional Neural Networks and Transfer Learning. In Proceedings of the 2023 2nd International Conference on Futuristic Technologies (INCOFT), Belagavi, Karnataka, India, 24–26 November 2023; pp. 1–6. [CrossRef]
50. Sun, Y.; Hao, Z.; Guo, Z.; Liu, Z.; Huang, J. Detection and Mapping of Chestnut Using Deep Learning from High-Resolution UAV-Based RGB Imagery. *Remote Sens.* **2023**, *15*, 4923. [CrossRef]
51. Kamal, M.; Phinn, S.; Johansen, K. Object-Based Approach for Multi-Scale Mangrove Composition Mapping Using Multi-Resolution Image Datasets. *Remote Sens.* **2015**, *7*, 4753–4783. [CrossRef]

52. Srivastava, P.K.; Malhi, R.K.M.; Pandey, P.C.; Anand, A.; Singh, P.; Pandey, M.K.; Gupta, A. Revisiting hyperspectral remote sensing: Origin, processing, applications and way forward. In *Hyperspectral Remote Sensing*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 3–21. [CrossRef]
53. Gallardo-Salazar, J.L.; Pompa-García, M. Detecting Individual Tree Attributes and Multispectral Indices Using Unmanned Aerial Vehicles: Applications in a Pine Clonal Orchard. *Remote Sens.* **2020**, *12*, 4144. [CrossRef]
54. Vannini, A.; Vettraino, A.M.; Belli, C.; Montaghi, A.; Natili, G. *New Technologies of Remote Sensing in Phytosanitary Monitoring: Application to Ink Disease of Chestnut [Castanea sativa Mill.; Latium]*; Italus Hortus: Sesto Fiorentino, Italy, 2005.
55. Pádua, L.; Marques, P.; Martins, L.; Sousa, A.; Peres, E.; Sousa, J.J. Monitoring of Chestnut Trees Using Machine Learning Techniques Applied to UAV-Based Multispectral Data. *Remote Sens.* **2020**, *12*, 3032. [CrossRef]
56. Pisner, D.A.; Schnyer, D.M. Support vector machine. In *Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 101–121. [CrossRef]
57. Hearst, M.; Dumais, S.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
58. Suthaharan, S. *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*; Integrated Series in Information Systems; Springer: Boston, MA, USA, 2016; Volume 36. [CrossRef]
59. Rish, I. An empirical study of the naive Bayes classifier. In Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, Seattle, WA, USA, 4 August 2001.
60. Xu, S. Bayesian Naïve Bayes classifiers to text classification. *J. Inf. Sci.* **2018**, *44*, 48–59. [CrossRef]
61. Kazakeviciute, A.; Olivo, M. A study of logistic classifier: Uniform consistency in finite-dimensional linear spaces. *J. Math. Oper. Res.* **2016**, *3*, 1–7.
62. Sebastiani, A.; Bertozzi, M.; Vannini, A.; Morales-Rodriguez, C.; Calfapietra, C.; Vaglio Laurin, G. Monitoring ink disease epidemics in chestnut and cork oak forests in central Italy with remote sensing. *Remote Sens. Appl. Soc. Environ.* **2024**, *36*, 101329. [CrossRef]
63. USDA-NRCS. *Keys to Soil Taxonomy*, 13th ed.; Technical Report; USDA Natural Resources Conservation Service: Washington, DC, USA, 2022.
64. IUSS Working Group WRB. *World Reference Base for Soil Resources, 2006: A Framework for International Classification, Correlation and Communication*; Technical Report; FAO: Rome, Italy, 2006.
65. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
66. Vannini, A.; Natili, G.; Anselmi, N.; Montaghi, A.; Vettraino, A.M. Distribution and gradient analysis of Ink disease in chestnut forests. *For. Pathol.* **2010**, *40*, 73–86. [CrossRef]
67. Maresi, G.; Turchetti, T. Management of diseases in chestnut orchards and stands: A significant prospect. *Adv. Hortic. Sci.* **2006**, *20*, 33–39.
68. Schomaker, M. *Crown-Condition Classification: A Guide to Data Collection and Analysis*; US Department of Agriculture, Forest Service, Southern Research Station: Asheville, NC, USA, 2007.
69. Bosshard, W. *Sanasilva—Kronenbilder*; Eidgenossische Anstalt fur das Forstliche Versuchswesen: Birmensdorf, Switzerland, 1986.
70. Topcon-Positioning-Systems Inc. *HiPer HR Operator's Manual*; Technical Report; Topcon Positioning Systems Inc.: Livermore, CA, USA, 2016.
71. Topcon-Positioning-Systems Inc. *MAGNET Field Layout Help*; Technical Report; Topcon Positioning Systems Inc.: Livermore, CA, USA, 2015.
72. Cutini, A. New management options in chestnut coppices: An evaluation on ecological bases. *For. Ecol. Manag.* **2001**, *141*, 165–174. [CrossRef]
73. Prada, M.; Cabo, C.; Hernández-Clemente, R.; Hornero, A.; Majada, J.; Martínez-Alonso, C. Assessing Canopy Responses to Thinnings for Sweet Chestnut Coppice with Time-Series Vegetation Indices Derived from Landsat-8 and Sentinel-2 Imagery. *Remote Sens.* **2020**, *12*, 3068. [CrossRef]
74. MicaSense-Inc. *RedEdge 3 User Manual MicaSense RedEdge 3 Multispectral Camera User Manual*; Technical Report; MicaSense-Inc.: Seattle, WA, USA, 2015.
75. Iglhaut, J.; Cabo, C.; Puliti, S.; Piermattei, L.; O'Connor, J.; Rosette, J. Structure from Motion Photogrammetry in Forestry: A Review. *Curr. For. Rep.* **2019**, *5*, 155–168. [CrossRef]
76. Aldrich, R.C. Detecting Disturbances in a Forest Environment. *Photogramm. Eng. Remote Sens.* **1975**, *41.1*, 39–48.
77. Coppin, P.; Jonckheere, I.; Nackaerts, K.; Muys, B.; Lambin, E. Review ArticleDigital change detection methods in ecosystem monitoring: A review. *Int. J. Remote Sens.* **2004**, *25*, 1565–1596. [CrossRef]

78. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W.; Freden, S.C.; Mercanti, E.P.; Becker, M.A. A Monitoring vegetation systems in the great plains with ERTS. In Proceedings of the Third ERTS Symposium, Washington DC, USA, 10–14 December 1973; Volume I, pp. 309–317. [CrossRef]
79. Woebbecke, D.; Meyer, G.; Von Bargen, K.; Mortensen, D. Shape features for identifying young weeds using image analysis. *Trans. ASAE* **1995**, *38*, 271–281. [CrossRef]
80. Gillies, S. Rasterio: Geospatial raster I/O for Python programmers. 2013. Available online: https://github.com/rasterio/rasterio (accessed on 4 March 2024).
81. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
82. ESRI, Redlands, CA: Environmental Systems Research Institute, ArcGIS Pro: Release 3.3.0. 2024.
83. Anaconda Software Distribution, Anaconda Documentation. 2020. Available online: https://docs.anaconda.com/ (accessed on 5 July 2023).
84. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, New York, NY, USA, 27–29 July 1992; COLT'92, pp. 144–152. [CrossRef]
85. Cortes, C.; Vapnik, V.; Saitta, L. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
86. Platt, J.C. *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods*; MIT Press: Cambridge, MA, USA, 1999.
87. Chang, C.C.; Lin, C.J. *LIBSVM: A Library for Support Vector Machines*; Association for Computing Machinery: New York, NY, USA, 2001.
88. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
89. Chan, T.F.; Golub, G.H.; Leveque, R.J. *Updating Formulae and a Pairwise Algorithm for Computing Sample Variances*; Association for Computing Machinery: New York, NY, USA, 1979. [CrossRef]
90. Raschka, S. MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. *J. Open Source Softw.* **2018**, *3*, 638. [CrossRef]
91. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: Experiences from the scikit-learn project. In Proceedings of the ECML PKDD Workshop: Languages for Data Mining and Machine Learning, Prague, Czech Republic, 23–27 September 2013; pp. 108–122. [CrossRef]
92. Pasta, S.; Sala, G.; La Mantia, T.; Bondì, C.; Tinner, W. The past distribution of Abies nebrodensis (Lojac.) Mattei: Results of a multidisciplinary study. *Veg. Hist. Archaeobotany* **2020**, *29*, 357–371. [CrossRef]
93. Padua, L.; Marques, P.; Martins, L.; Sousa, A.; Peres, E.; Sousa, J.J. Estimation of Leaf Area Index in Chestnut Trees using Multispectral Data from an Unmanned Aerial Vehicle. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 6503–6506. [CrossRef]
94. Pádua, L.; Chiroque-Solano, P.M.; Marques, P.; Sousa, J.J.; Peres, E. Mapping the Leaf Area Index of Castanea sativa Miller Using UAV-Based Multispectral and Geometrical Data. *Drones* **2022**, *6*, 422. [CrossRef]
95. Mahlein, A.K.; Steiner, U.; Dehne, H.W.; Oerke, E.C. Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precis. Agric.* **2010**, *11*, 413–431. [CrossRef]
96. Pinter, P.J., Jr.; Hatfield, J.L.; Schepers, J.S.; Barnes, E.M.; Moran, M.S.; Daughtry, C.S.; Upchurch, D.R. Remote Sensing for Crop Management. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 647–664. [CrossRef]
97. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sens.* **2021**, *13*, 586. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.