horticulturae

MDPI

Article

Detection of *Eulecanium giganteum* Infestation in Jujube Orchards Based on UAV Multispectral Remote Sensing

Yueteng Chao [1,2,3,†], Yulong Wang [1,2,3,†], Zhuoyang Xu [1,2,3], Ruohong He [1,2,3], Rina Yangdao [1,2,3], Hongqiang Dong [4] and Ping Li [1,2,3,*]

[1] College of Mechanical and Electrical Engineering, Tarim University, Alar 843300, China
[2] Xinjiang Production and Construction Corps (XPCC) Key Laboratory of Utilization and Equipment of Special Agricultural and Forestry Products in Southern Xinjiang, Alar 843300, China
[3] Modern Agricultural Engineering Key Laboratory at Universities of Education Department of Xinjiang Uygur Autonomous Region, Tarim University, Alar 843300, China
[4] College of Agronomy, Tarim University, Alar 843300, China
* Correspondence: 119980015@taru.edu.cn
† These authors contributed equally to this work.

Abstract

The application of remote sensing technology to monitor agricultural pests and diseases provides an effective approach for rapid and non-destructive assessment of crop health. This study evaluated a UAV multispectral remote sensing framework integrating vegetation indices and SPAD information for assessing *Eulecanium giganteum* infestation in jujube orchards. Initially, 24 vegetation indices were extracted from canopy multispectral imagery, and Pearson correlation analysis was used for feature screening, resulting in a refined subset of 19 vegetation indices. Seven machine learning algorithms were evaluated during model development, with Grid Search combined with tree-level grouped 7-fold cross-validation for hyperparameter selection. Vegetation indices were used as predictor variables, while SPAD value and infestation severity were treated as response variables in the regression and classification tasks, respectively. During internal model development, the CatBoost-19opt model achieved an R^2 of 0.81 and an RMSE of 1.84 for SPAD estimation. On the independent test set consisting of 58 jujube trees, the previously selected model achieved an R^2 of 0.84 and an RMSE of 1.56. During internal model development, incorporating measured SPAD values improved pest classification performance, with XGBoost-19sopt achieving an OA of 0.89 and a Kappa coefficient of 0.84. On the independent test set, the measured-SPAD configuration achieved an OA of 0.88, whereas both the VI-only and VI combined with predicted-SPAD configurations achieved an OA of 0.71. Thus, incorporating predicted SPAD did not improve overall classification accuracy relative to using vegetation indices alone. Further model optimization and additional validation across different orchards, growing seasons, and phenological stages are required before the approach can be considered for operational large-scale monitoring.

Keywords: SPAD; vegetation indices; machine learning; chlorophyll estimation; multi-source feature fusion

Check for updates

Academic Editor: Arturo Alvino

Received: 4 September 2026
Revised: 11 October 2026
Accepted: 13 October 2026
Published: 16 October 2026

Copyright: © 2026 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.

1. Introduction

Xinjiang is currently China's largest jujube-producing region, responsible for over 90% of the nation's annual yield [1]. However, orchards in southern Xinjiang face significant pest and disease threats, particularly widespread infestations of the jujube giant scale. This pest

damages the trees by extracting sap from leaves and branches, which reduces their ability to photosynthesize. As a result, leaves turn yellow and fruit shrivels; severe infestations can lead to the death of branches or the entire tree. Therefore, rapid and accurate monitoring of *Eulecanium giganteum* infestation is important for reducing yield loss and supporting timely pest management in jujube orchards.

Traditional monitoring relies on manual field surveys, which are labor-intensive, time-consuming, and impractical for evaluating large areas simultaneously [2,3]. Fortunately, recent advancements in remote sensing technology now permit the non-destructive monitoring of pests and diseases across expansive orchards. Specifically, unmanned aerial vehicles (UAV) equipped with multispectral sensors offer a flexible, cost-effective, and high-resolution method for efficiently gathering agricultural data [4,5].

Pest and disease stress alters plant cellular structure, degrades pigments, and causes water loss, thereby changing canopy spectral reflectance. This physiological response forms the foundation for monitoring crop health via remote sensing [6]. In the case of *Eulecanium giganteum*, the insect feeds on jujube leaves and branches by sucking sap, which can weaken nutrient and water transport, accelerate leaf yellowing, and reduce chlorophyll content. Since SPAD values are closely related to leaf chlorophyll content, they can be used as a physiological indicator of pest-induced stress. Meanwhile, changes in chlorophyll content and canopy condition may affect reflectance in the green, red, red-edge, and near-infrared bands, providing a theoretical basis for using multispectral vegetation indices to distinguish different infestation levels.

Today, remote sensing applications have expanded from diagnosing individual plants to assessing regional health on a large scale. Because traditional linear models struggle with the high dimensionality and complex, non-linear features of UAV data, researchers increasingly rely on machine learning and ensemble algorithms [7,8]. For example, Zhuang et al. demonstrated that ensemble models like XGBoost outperform Support Vector Machines and Partial Least Squares Regression when estimating chlorophyll content in woody plants with complex canopies and waxy leaves [9]. Similarly, Huang et al. successfully applied Gaussian Process Regression to estimate canopy chlorophyll across various fruit tree species and growth stages [10]. To further optimize these models, researchers frequently use statistical methods like the Successive Projections Algorithm and Analysis of Variance to isolate key spectral bands and construct new vegetation indices. For instance, Zhang et al. identified NDSI and SAVI as the most sensitive indices for monitoring the severity of larch pest infestations [11]. Building on this approach, Yu et al. developed a new vegetation index derived from sensitive spectral bands to detect mirid bug stress in cotton, improving identification accuracy [12].

Despite these advancements, models relying exclusively on single-source spectral data remain susceptible to interference from soil backgrounds and complex canopy structures. Consequently, fusing multiple data features has emerged as a core strategy to improve model accuracy [13,14]. For example, Ai et al. demonstrated that combining vegetation indices with texture features improved chlorophyll estimation in winter wheat by over 20% compared to using isolated features [15]. Building on this concept, Wang et al. incorporated spectral, textural, and structural features to estimate chlorophyll content in cotton, using SHAP analysis to map how these contributions vary across different flight altitudes [16]. This approach is particularly valuable for complex targets; Yan et al. fused multispectral indices with canopy texture features to achieve more accurate leaf chlorophyll estimates in fragrant pear trees [17]. Beyond extracting multiple features from a single source, integrating data across entirely different sensors has also shown potential for practical application. For instance, Sa et al. integrated UAV multispectral and LiDAR data into an SVM model to classify larch caterpillar damage, achieving an Overall Accuracy of

95.8% [18]. Similarly, Hornero et al. paired multispectral data with thermal imaging to enable the early detection of oak decline [19].

Narrowing the focus to jujube trees, remote sensing technology has recently made progress in both estimating physiological parameters and monitoring crop health. In terms of physiological parameters, Wu et al. developed a hybrid machine learning approach integrating hyperspectral data and image color channels to estimate chlorophyll content in jujube leaves [20]. Nigela et al. accurately estimated jujube chlorophyll content by applying dimensionality reduction and Support Vector Regression to ground-based hyperspectral data [21]. Similarly, Wang et al. developed a Back-Propagation Neural Network model to predict leaf chlorophyll (SPAD) values using near-infrared spectroscopy [22]. Shifting to pest and disease monitoring, Liu et al. used a near-infrared hyperspectral imaging system to model external pest damage on jujube fruits, achieving a 93.1% recognition rate [23]. Furthermore, Jiang et al. successfully employed similar hyperspectral technology to differentiate between the various infection stages of black spot disease in winter jujubes [24].

Although remote sensing studies on jujube have made progress in physiological parameter estimation and pest or disease identification, most existing research still relies heavily on proximal hyperspectral or near-infrared spectroscopy and is mainly conducted at the leaf or fruit scale. Therefore, these approaches remain limited for canopy-scale and large-area monitoring of *Eulecanium giganteum* infestation in jujube orchards. To address this gap, this study developed a UAV multispectral remote sensing framework that integrates canopy vegetation indices with SPAD information for pest infestation assessment. In this framework, vegetation indices extracted from UAV multispectral imagery were used as predictor variables, while SPAD value and infestation severity were treated as response variables in the regression and classification tasks, respectively. The coefficient of determination (R^2) and root mean square error (RMSE) were used to evaluate SPAD regression models, whereas Overall Accuracy, Kappa coefficient, and F1-score were used only as evaluation and model selection metrics for pest classification models. Specifically, this study aimed to: (1) extract vegetation indices from UAV multispectral imagery and analyze their relationships with SPAD values; (2) compare multiple machine learning algorithms for SPAD estimation; (3) construct direct pest detection models using vegetation indices and measured SPAD values; and (4) evaluate the feasibility of indirect pest detection using predicted SPAD values. This study provides a canopy-scale modeling strategy for monitoring jujube trees under *E. giganteum* stress and offers a basis for further development of UAV-based orchard pest monitoring methods.

2. Materials and Methods

2.1. Study Area

The study area is located in Alar City, Xinjiang (40°32′ N, 81°17′ E), encompassing the jujube orchards of Tarim University and adjacent farms (Figure 1). encompassing the jujube orchards of Tarim University and adjacent farms. Situated at the northwestern margin of the Tarim Basin, the region features a typical warm-temperate, extreme continental arid climate. It is characterized by an average annual temperature of approximately 11.2 °C, high diurnal temperature variations, abundant sunshine, and low annual precipitation (40–50 mm). These abundant light and thermal resources create an optimal environment for jujube cultivation. Within this experimental area, 40 individual jujube trees were selected as independent sampling points. The selected trees covered the observed infestation gradient in the orchard, including healthy, mildly infested, and severely infested trees, to improve the representativeness of the dataset. Field data were collected at approximately two-week intervals from 15 May to 15 July 2025, predominantly covering the flowering and fruit-

setting stage of jujube trees. Each sampled tree on each sampling date was treated as one tree-level sample, resulting in a total of 200 tree-level samples for subsequent analysis.

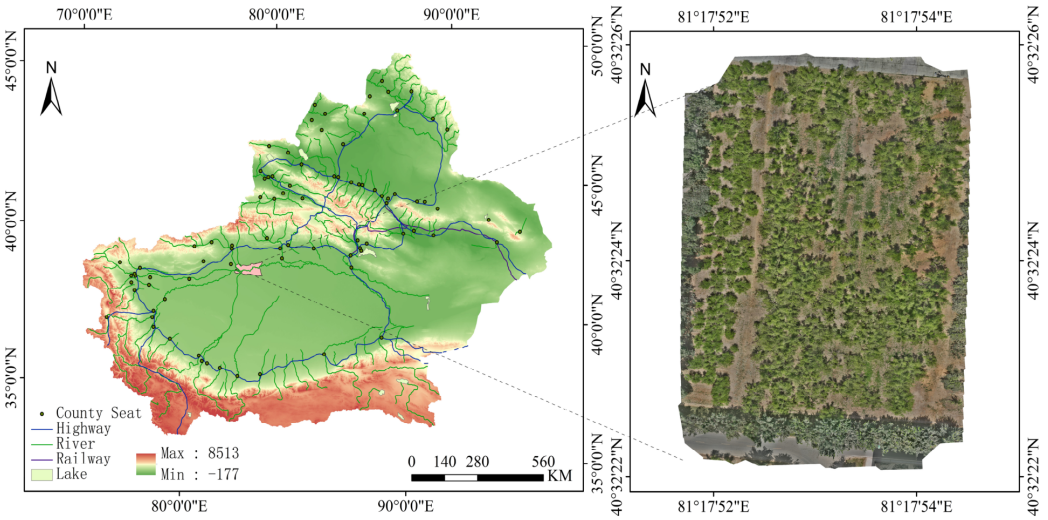


Figure 1. Study Area.

2.2. Data Acquisition

2.2.1. Ground Truth Data Collection

First, the precise geographic coordinates of each sampled tree were recorded. Specifically, an A100 RTK GNSS receiver (Qianxun Location Network (Zhejiang) Co., Ltd., Huzhou, China) was used to log these waypoints into a database for subsequent extraction and analysis. To improve the representativeness of leaf sampling, leaves were collected using a stratified sampling method. Each tree canopy was divided into four cardinal directions (east, south, west, and north). From each direction, four fresh leaves were randomly selected across the upper, middle, and lower canopy strata, resulting in a total of 16 leaves per tree. Immediately after collection, the samples were stored in labeled, opaque, sealed bags to prevent moisture loss and chlorophyll degradation. Finally, they were promptly transported to the laboratory for analysis.

A SPAD-502Plus portable chlorophyll meter (Konica Minolta, Inc., Osaka, Japan)was used to measure leaf SPAD values, providing a quantitative indicator of chlorophyll content [25]. Following instrument calibration, measurements were taken at three evenly spaced points (upper, middle, and lower) along each leaf. Care was taken to strictly avoid the midrib to prevent structural interference. The mean of the three readings was recorded as the SPAD value for each leaf. Subsequently, the SPAD values of the 16 sampled leaves from each tree were averaged, and this mean value was used as the representative tree-level SPAD value for subsequent model construction and analysis.

Before UAV multispectral data acquisition, an independent field infestation survey was conducted for each sampled tree with the participation of researchers from the College of Agronomy. The canopy of each tree was divided into four cardinal directions, and 25 leaves were examined in each direction, resulting in a total of 100 leaves per tree. *Eulecanium giganteum* infestation was confirmed by direct observation of scale insects, scale covers, or characteristic attachment features on the sampled leaves and associated branches, rather than solely based on nonspecific symptoms such as leaf yellowing. The number of damaged leaves among the 100 examined leaves was recorded and used to calculate the leaf infestation rate. Because the survey was completed before UAV image acquisition, the assessors had no access to remote-sensing information or subsequent model outputs when assigning infestation severity classes.

The field data collection procedure is illustrated in Figure 2.

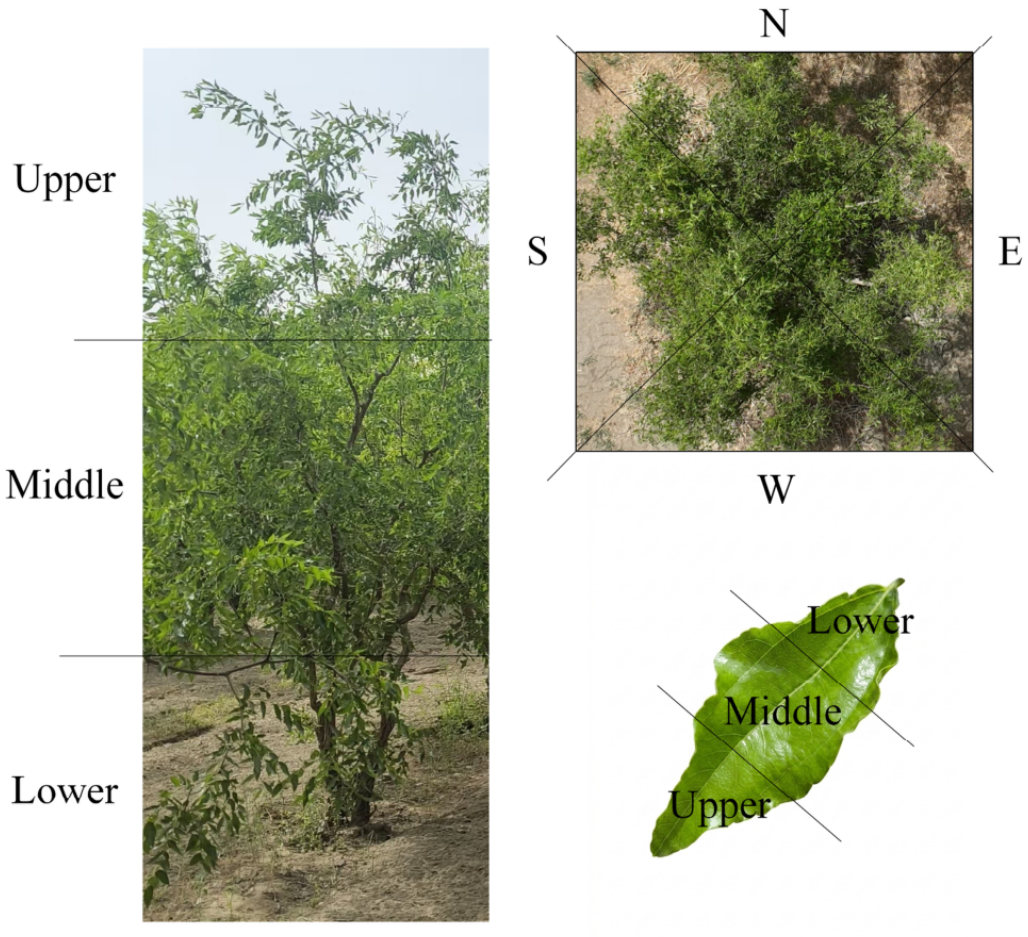


Figure 2. Example of data collection.

Infestation severity was classified with reference to the grading criteria for *Eulecanium giganteum* specified in LY/T 1681-2006, Standard of Forest Pests Occurrence and Disaster [26]. According to this standard, leaf infestation rate is defined as the percentage of damaged leaves relative to the total number of leaves investigated:

$$\mathrm{LIR} = \frac{N_d}{N_t} \times 100\% \quad (1)$$

where LIR is the leaf infestation rate; N_d is the number of damaged leaves; and N_t is the total number of leaves investigated.

For the three-class classification task in this study, the original grading scheme was adapted to distinguish healthy trees from trees exhibiting different levels of visible infestation. Trees with no damaged leaves were classified as Healthy (0%), trees with a leaf infestation rate of 1–10% were classified as Mild, and trees with a leaf infestation rate of ≥11% were classified as Severe (Table 1).

Table 1. Classification Criteria.

Grade	Classification Criteria	Scale Value
Healthy	Leaf infestation rate = 0%	0
Mild	Leaf infestation rate = 1–10%	1
Severe	Leaf infestation rate ≥ 11%	2

2.2.2. Multispectral Data Acquisition

UAV imagery was acquired using a DJI Mavic 3 Multispectral drone (DJI, Shenzhen, China), equipped with a visible light sensor and four multispectral bands: Green (G), Red (R), Red Edge (RE), and Near-Infrared (NIR). To minimize the effects of shadows and varying illumination, data was collected under clear skies between 12:00 and 16:00. To capture high-resolution spectral images of the canopy, the flight altitude was set to 20 m with an 80% forward overlap and a 70% side overlap. Comprehensive flight parameters are detailed in Table 2. Additionally, to account for atmospheric and illumination variables, a standard radiometric calibration panel was photographed orthogonally 2 m above ground prior to each flight. This provided the necessary baseline data for subsequent radiometric correction.

Table 2. Technical parameters of UAV flight and multispectral sensor.

Technical Parameter	Value
Flight Height	20 m
Forward Overlap	80%
Side Overlap	70%
Flight Speed	1.2 m/s
Maximum Image Size	2592 × 1944 pixels
Green (G)	560 nm
Red (R)	650 nm
Red Edge (RE)	730 nm
Near-Infrared (NIR)	840 nm

2.3. Data Processing

2.3.1. Data Preprocessing

Following acquisition, the raw multispectral UAV images required several preprocessing steps before analysis. First, the DJI Smart Agriculture platform was used to stitch and crop the data, generating a comprehensive multispectral orthomosaic of the study area in TIFF format. Next, the field coordinates were processed and formatted using ArcMap 10.8 (Esri, Redlands, CA, USA) software. Finally, these coordinates were overlaid onto the TIFF image within ENVI 5.6 software (NV5 Geospatial Solutions, Inc., Broomfield, CO, USA) to accurately map the geographic locations of all samples. The final multispectral orthomosaic had a spatial resolution of approximately 0.58 cm/pixel.

Regions of Interest (ROIs) were delineated around the geographic coordinates of each sampled tree to extract canopy-level Digital Number (DN) values. To reduce the influence of single-pixel noise and local canopy heterogeneity on spectral feature extraction, the DN values of all valid pixels within each ROI were averaged and used as the representative DN value for the corresponding sample. Additionally, to estimate surface spectral reflectance, the acquired images were radiometrically calibrated using a 50% reflectance panel. This calibration was performed using the following equation [27]:

$$R_{target} = DN_{target} \times \frac{R_{panel}}{DN_{panel}} \quad (2)$$

where R_{target} is the corrected surface reflectance of the target feature; DN_{target} is the raw digital number extracted for the target; R_{panel} is the known standard reflectance of the calibration panel; and DN_{panel} is the average DN value of the calibration panel within the image.

2.3.2. Calculation of Vegetation Indices

Because single-band reflectance is often insufficient for characterizing complex physiological and biochemical changes, this study constructed 24 vegetation indices to maximize the utility of the multispectral data and improve sensitivity to pest stress. As detailed in Table 3, these indices were derived from four UAV-acquired spectral bands: Green, Red, Red Edge, and Near-Infrared.

Table 3. The 24 vegetation indices.

VIS	Formula	Reference
NDVI	$(\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R)$	[27]
RVI	ρ_{NIR}/ρ_R	[28]
GNDVI	$(\rho_{NIR} - \rho_G)/(\rho_{NIR} + \rho_G)$	[29]
OSAVI	$(\rho_{NIR} - \rho_R)/(\rho_{NIR} + \rho_R + 0.16)$	[30]
MSAVI	$0.5\left[2\rho_{NIR} + 1 - \sqrt{(2\rho_{NIR} + 1)^2 - 8(\rho_{NIR} - \rho_R)}\right]$	[31]
MSR	$(\rho_{NIR}/\rho_R - 1)/\sqrt{\rho_{NIR}/\rho_R + 1}$	[31]
EVI2	$2.5(\rho_{NIR} - \rho_R)/(\rho_{NIR} + 2.4\rho_R + 1)$	[32]
RDVI	$(\rho_{NIR} - \rho_R)/\sqrt{\rho_{NIR} + \rho_R}$	[33]
DVI	$\rho_{NIR} - \rho_R$	[34]
IPVI	$\rho_{NIR}/(\rho_{NIR} + \rho_R)$	[33]
NLI	$(\rho_{NIR}^2 - \rho_R)/(\rho_{NIR}^2 + \rho_R)$	[35]
TVI	$0.5[120(\rho_{NIR} - \rho_G) - 200(\rho_R - \rho_G)]$	[36]
GCI	$(\rho_{NIR}/\rho_G) - 1$	[37]
MTVI	$1.2[1.2(\rho_{NIR} - \rho_G) - 2.5(\rho_R - \rho_G)]$	[38]
NDRE	$(\rho_{NIR} - \rho_{RE})/(\rho_{NIR} + \rho_{RE})$	[39]
MCARI2	$[(\rho_{RE} - \rho_R) - 0.2(\rho_{RE} - \rho_G)](\rho_{RE} - \rho_R)/(\rho_{RE}/\rho_R)$	[40]
RECI	$(\rho_{NIR}/\rho_{RE}) - 1$	[40]
MTCI	$(\rho_{NIR} - \rho_{RE})/(\rho_{RE} - \rho_R)$	[41]
MCARI	$[(\rho_{RE} - \rho_R) - 0.2(\rho_{RE} - \rho_G)](\rho_{RE} - \rho_R)$	[42]
TCARI	$3[(\rho_{RE} - \rho_R) - 0.2(\rho_{RE} - \rho_G)(\rho_{RE}/\rho_R)]$	[43]
CSI	$TCARI/OSAVI$	[44]
MRVI	$(\rho_{RE}/\rho_R - 1)/(\sqrt{\rho_{NIR}/\rho_R} - 1)$	[45]
RENDVI	$(\rho_{NIR} - \rho_{RE})/(\rho_{NIR} + \rho_{RE})$	[46]
RESR	ρ_{RE}/ρ_R	[46]

2.4. Feature Dimensionality Reduction

Pearson correlation analysis was used as an exploratory screening procedure to evaluate the association between each vegetation index and SPAD. Vegetation indices showing relatively weak correlations with SPAD ($|r| < 0.30$) were excluded from the reduced feature set. This procedure was used to obtain a more compact predictor set for subsequent model development and is therefore referred to as SPAD-correlation screening rather than a redundancy-based feature-selection procedure [47].

2.5. Parameter Optimization

The performance of machine learning models is influenced by their hyperparameter configurations. During model development, Grid Search was employed to identify suitable hyperparameter combinations for the multi-dimensional feature models. By pre-defining candidate ranges and step sizes, Grid Search systematically evaluated the specified parameter combinations based on the internal cross-validation results [48]. For SPAD regression, two feature sets were considered: the original 24 vegetation indices (24VI) and the reduced set of 19 vegetation indices (19VI). For infestation severity classification, four feature configurations were evaluated: 24VI, 19VI, 24VI combined with measured SPAD, and 19VI

combined with measured SPAD. These feature configurations were used to compare the performance of the seven machine learning algorithms.

To improve the transparency and reproducibility of the modelling process, the hyperparameter settings of the seven machine learning algorithms were further documented. Table 4 summarizes the hyperparameter search ranges, step sizes, and optimized parameter configurations for the models using the 19VI feature set [49,50].

Table 4. Hyperparameter search ranges, step sizes, and optimized parameter configurations for the 19VI regression models.

Model	Key Hyperparameter	Search Range	Optimal Value
Ridge	alpha	0.0001–0.01, step = 0.0001	alpha = 0.001
KNN	n_neighbors	1–15, step = 1	n_neighbors = 3
SVM	C, gamma, kernel	C: 1–20, step = 1; gamma: 0.1–5, step = 0.1; kernel: linear, poly, rbf, sigmoid	C = 10; gamma = 1; kernel = rbf
BPNN	hidden_layer_sizes, max_iter	First hidden layer: 20–200, step = 10; second hidden layer: 20–120, step = 10; max_iter: 200–1000, step = 10	hidden_layer_sizes = (140, 100); max_iter = 400
RF	n_estimators, max_depth	n_estimators: 100–1000, step = 10; max_depth: 10–50, step = 1	n_estimators = 950; max_depth = 40
XGBoost	learning_rate, n_estimators, max_depth	learning_rate: 0.01–0.20, step = 0.01; n_estimators: 10–200, step = 10; max_depth: 1–10, step = 1	learning_rate = 0.07; n_estimators = 100; max_depth = 9
CatBoost	iterations, learning rate, depth	iterations: 1000–5000, step = 100; learning_rate: 0.01–0.11, step = 0.01; depth: 1–15, step = 1	iterations = 2000; learning_rate = 0.11; depth = 12

The models were implemented in Python. Ridge regression, KNN, SVR, BPNN, and RF were developed using the scikit-learn library, while XGBoost and CatBoost were implemented using their respective Python libraries. The computational environment inspected for reproducibility included Python 3.8.19, NumPy 1.24.3, pandas 2.0.3, SciPy 1.10.1, scikit-learn 1.3.0, XGBoost 2.1.1, and CatBoost 1.2.3.

Regarding model configuration, KNN employed distance-based weighting (weights = 'distance'), while the optimized SVM model used a radial basis function (RBF) kernel. The CatBoost implementation specified verbose = False and allow_writing_files = False. Parameters not explicitly specified retained their respective library default settings. A random_state of 42 was explicitly specified for BPNN, RF, and XGBoost.

2.6. Evaluation Metrics

To comprehensively evaluate the SPAD regression models, the coefficient of determination $\left(R^2\right)$ and root mean square error (RMSE) were employed to assess model precision and accuracy. The R^2 metric indicates the goodness of fit, with values closer to 1 denoting superior model performance. Conversely, RMSE measures the deviation between predicted and observed values, where smaller values indicate higher predictive accuracy [51].

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{4}$$

where n is the total number of samples; y_i is the measured SPAD value of the i-th sample; $\hat{y}_i$ represents the predicted value from the model; and $\overline{y}$ is the mean of all measured values.

To evaluate the performance of the infestation detection models, Overall Accuracy (OA), the F1-score (F1), and the Kappa coefficient were utilized. OA quantifies the model's overall predictive capability. The F1-score is the harmonic mean of precision and recall and summarizes classification performance. Higher values indicate better combined performance in precision and recall. Finally, the Kappa coefficient adjusts for chance agreement, providing a more reliable assessment of true classification accuracy [52,53].

$$OA = \frac{\sum_{i=1}^{k} x_{ii}}{N} \tag{5}$$

$$F1 = \frac{1}{k}\sum_{i=1}^{k}\frac{2 \times x_{ii}}{x_{+i} + x_{i+}} \tag{6}$$

$$Kappa = \frac{N\sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k}\left(x_{i+} \times x_{+i}\right)}{N^2 - \sum_{i=1}^{k}\left(x_{i+} \times x_{+i}\right)} \tag{7}$$

where N is the total number of samples; k denotes the total number of classes; x_{ii} represents the number of correctly classified samples for the i-th class; and x_{i+} and x_{+i} represent the marginal totals of the actual and predicted samples for class i, respectively.

To account for the repeated-measures structure of the dataset and reduce tree-level pseudoreplication, a tree-level grouped 7-fold cross-validation strategy was employed during model development. All observations originating from the same sampled tree were assigned to the same fold, thereby preventing observations from the same tree from appearing simultaneously in the training and validation folds. In each iteration, six folds were used for model training and the remaining fold was used for validation. In addition, multiple valid pixels within each tree-level ROI were averaged to obtain a single representative spectral value for the corresponding tree-date observation, and individual pixels were not treated as independent machine-learning samples. The resulting cross-validation metrics were used for internal model comparison and hyperparameter selection and should not be interpreted as fully independent estimates of model generalization performance. The complete research workflow is illustrated in Figure 3.

In addition to the grouped cross-validation used during model development, an independent test set consisting of 58 jujube trees was collected on 10 July 2025 from a separate farmer-managed jujube orchard. At the time of data acquisition, the test trees were in the flowering and fruit-setting stage. The same UAV acquisition settings and radiometric-calibration procedures as those used for the development dataset were applied. SPAD values and infestation severity were measured following the same procedures used for the development dataset. Each test tree contributed a single independent observation, and no repeated observations were included in the independent test set. These trees were completely excluded from feature screening, model training, hyperparameter optimization, cross-validation, and model selection. The independent test set contained 18 Healthy, 21 Mild, and 19 Severe trees. The independent test set was used only for the final evaluation of SPAD regression models and the comparison of three infestation-classification input configurations: 19VI alone, 19VI combined with measured SPAD, and 19VI combined with predicted SPAD.

Therefore, the independent evaluation represents spatial validation across separate orchards under overlapping seasonal and phenological conditions, rather than validation across different growing seasons or phenological stages.

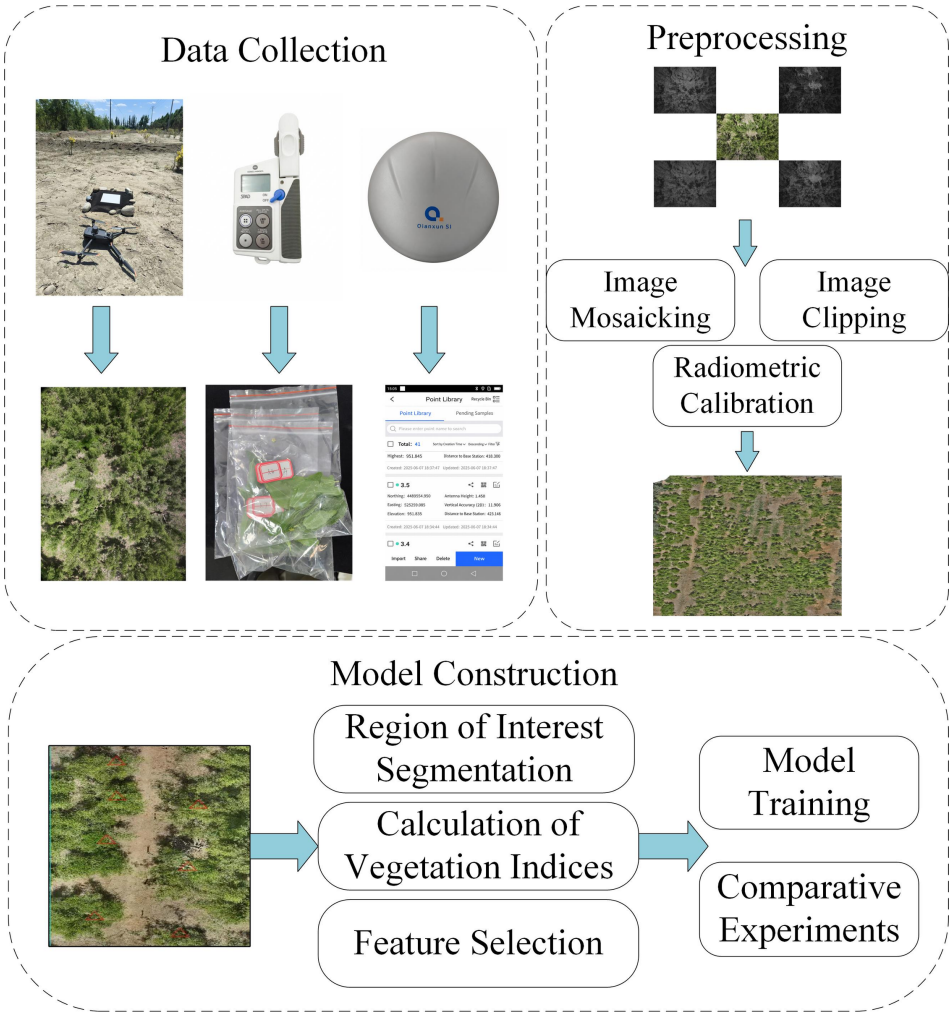


Figure 3. Data acquisition and processing.

3. Results

3.1. Correlation Between Vegetation Indices and Chlorophyll Content

Pearson correlation analysis was conducted to evaluate the relationship between the 24 vegetation indices and the jujube tree SPAD values (Figure 4). Key features were identified by evaluating the correlation coefficients (r) and significance levels (p-value) between each index and the target SPAD measurements. The results indicated that all vegetation indices (VIs) were significantly correlated with SPAD ($p \leq 0.01$). While most VIs exhibited positive correlations, CSI was the sole index to demonstrate a negative relationship. Specifically, NLI achieved the highest correlation (r = 0.55), followed by MTVI and TVI (r = 0.54). According to the SPAD-correlation screening criterion defined in Section 2.4, five vegetation indices (MCARI2, MTCI, TCARI, CSI, and MRVI) were excluded, resulting in a reduced set of 19 vegetation indices for subsequent analysis.

3.2. Model Performance Evaluation

3.2.1. SPAD Regression Model Performance

The internal model-development results for SPAD regression using the 24VI and 19VI feature sets are presented in Table 5. Models with optimized hyperparameters are denoted by the suffix "-opt". The reported R^2 and RMSE values were used for internal model comparison, rather than as independent estimates of generalization performance.

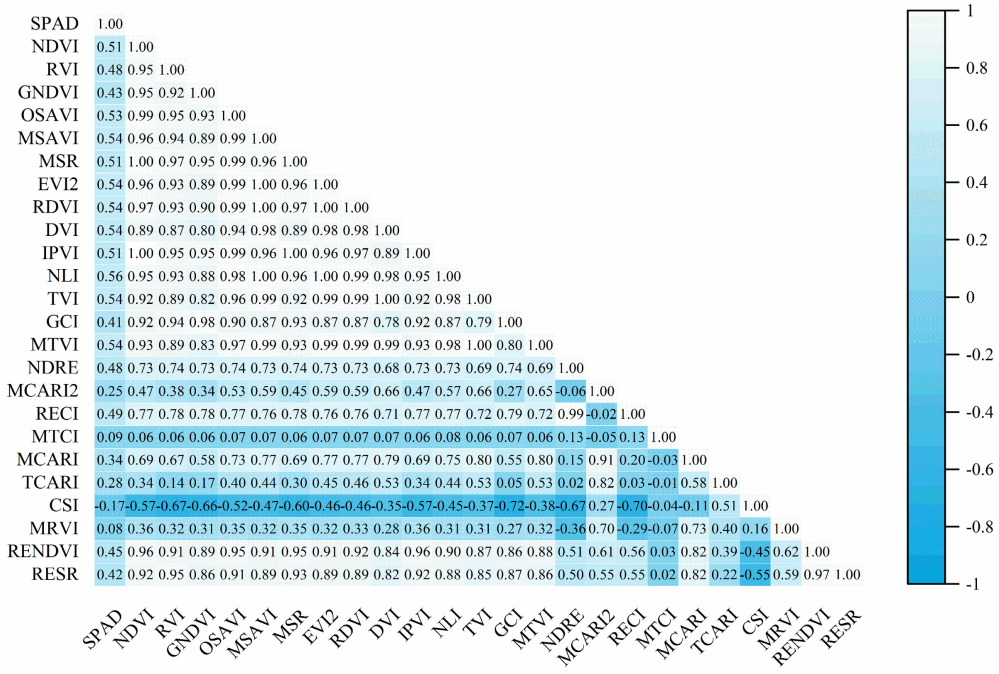


Figure 4. Heatmaps of correlation coefficients.

Table 5. Results of regression models.

Features	Model	R^2	RMSE	Model	R^2	RMSE
24VI	BPNN	0.46	3.18	BPNN-opt	0.47	3.17
	KNN	0.79	1.98	KNN-opt	0.80	1.89
	SVM	0.29	3.67	SVM-opt	0.61	2.72
	RF	0.75	2.16	RF-opt	0.75	2.15
	XGBoost	0.60	2.74	XGBoost-opt	0.70	2.36
	CatBoost	0.77	2.04	CatBoost-opt	0.81	1.88
	Ridge	0.40	3.37	Ridge-opt	0.44	3.25
19VI	BPNN	0.40	3.37	BPNN-opt	0.49	3.17
	KNN	0.79	1.98	KNN-opt	0.80	1.92
	SVM	0.28	3.71	SVM-opt	0.52	3.01
	RF	0.74	2.17	RF-opt	0.75	2.17
	XGBoost	0.58	2.79	XGBoost-opt	0.71	2.33
	CatBoost	0.77	2.07	CatBoost-opt	0.81	1.84
	Ridge	0.36	3.48	Ridge-opt	0.41	3.33

3.2.2. Pest Infestation Classification Performance

The internal classification results for the four feature configurations (24VI, 19VI, 24VI combined with measured SPAD, and 19VI combined with measured SPAD) are summarized in Tables 6 and 7. Table 6 presents the results obtained using vegetation indices alone, whereas Table 7 presents the results obtained by incorporating measured SPAD values. The reported OA, Kappa, and F1 values were used for internal model comparison.

Following the internal model comparisons presented in Tables 6 and 7, an independent test-set evaluation was conducted to further assess the classification performance of the selected models. Three input configurations were compared: 19VI alone, 19VI combined with measured SPAD, and 19VI combined with predicted SPAD. This comparison was designed to evaluate the contribution of measured SPAD information and determine whether incorporating predicted SPAD could improve classification performance relative

to using vegetation indices alone. The overall workflow of model development and independent validation is illustrated in Figure 5.

Table 6. Classification performance of pest infestation levels based on single spectral features.

Features	Model	OA	Kappa	F1	Model	OA	Kappa	F1
24VI	BPNN-24	0.57	0.33	0.55	BPNN-24opt	0.60	0.41	0.58
	KNN-24	0.69	0.53	0.68	KNN-24opt	0.74	0.61	0.74
	SVM-24	0.58	0.38	0.56	SVM-24opt	0.78	0.67	0.78
	RF-24	0.76	0.64	0.76	RF-24opt	0.77	0.65	0.77
	XGBoost-24	0.64	0.47	0.64	XGBoost-24opt	0.76	0.64	0.76
	CatBoost-24	0.72	0.58	0.72	CatBoost-24opt	0.77	0.66	0.77
	Ridge-24	0.56	0.35	0.55	Ridge-24opt	0.58	0.37	0.56
19VI	BPNN-19	0.56	0.34	0.54	BPNN-19opt	0.59	0.38	0.57
	KNN-19	0.69	0.53	0.69	KNN-19opt	0.73	0.60	0.73
	SVM-19	0.58	0.37	0.56	SVM-19opt	0.79	0.68	0.79
	RF-19	0.75	0.63	0.75	RF-19opt	0.76	0.64	0.75
	XGBoost-19	0.64	0.46	0.63	XGBoost-19opt	0.74	0.62	0.74
	CatBoost-19	0.71	0.57	0.71	CatBoost-19opt	0.77	0.65	0.77
	Ridge-19	0.56	0.34	0.54	Ridge-19opt	0.56	0.35	0.55

Table 7. Classification results of models based on multi-source feature fusion.

Features	Model	OA	Kappa	F1	Model	OA	Kappa	F1
24VI + SPAD	BPNN-24s	0.65	0.47	0.64	BPNN-24sopt	0.67	0.51	0.67
	KNN-24s	0.77	0.65	0.77	KNN-24sopt	0.82	0.74	0.82
	SVM-24s	0.63	0.45	0.62	SVM-24sopt	0.87	0.81	0.87
	RF-24s	0.86	0.79	0.86	RF-24sopt	0.86	0.79	0.86
	XGBoost-24s	0.78	0.67	0.78	XGBoost-24sopt	0.89	0.83	0.89
	CatBoost-24s	0.84	0.76	0.84	CatBoost-24sopt	0.88	0.82	0.88
	Ridge-24s	0.62	0.43	0.61	Ridge-24sopt	0.62	0.43	0.61
19VI + SPAD	BPNN-19s	0.62	0.43	0.61	BPNN-19sopt	0.65	0.48	0.64
	KNN-19s	0.78	0.67	0.78	KNN-19sopt	0.82	0.73	0.82
	SVM-19s	0.63	0.45	0.61	SVM-19sopt	0.88	0.82	0.88
	RF-19s	0.86	0.79	0.86	RF-19sopt	0.86	0.79	0.86
	XGBoost-19s	0.77	0.66	0.77	XGBoost-19sopt	0.89	0.84	0.89
	CatBoost-19s	0.77	0.65	0.77	CatBoost-19sopt	0.88	0.83	0.88
	Ridge-19s	0.62	0.43	0.61	Ridge-19sopt	0.62	0.43	0.61

3.3. Comparative Analysis of Models

3.3.1. Comparative Analysis of SPAD Regression Models

Using the complete 24VI dataset, CatBoost, KNN, and RF demonstrated robust performance under default parameters, with KNN achieving the highest initial accuracy (R^2 = 0.79, RMSE = 1.98). Hyperparameter tuning via Grid Search improved the internal performance of several models, although the magnitude of improvement varied across algorithms. CatBoost-opt emerged as the top performer (R^2 = 0.81, RMSE = 1.88), achieving a 5% increase in R^2 and a 7% reduction in RMSE. However, algorithm sensitivity to dimensionality reduction varied across algorithms. Linear and traditional models, such as Ridge and SVM, relied heavily on feature completeness. For instance, the R^2 of the SVM model dropped from 0.61 to 0.52 following feature pruning. Conversely, non-linear models (CatBoost, KNN, and RF) remained relatively unaffected by the reduced feature dimensionality.

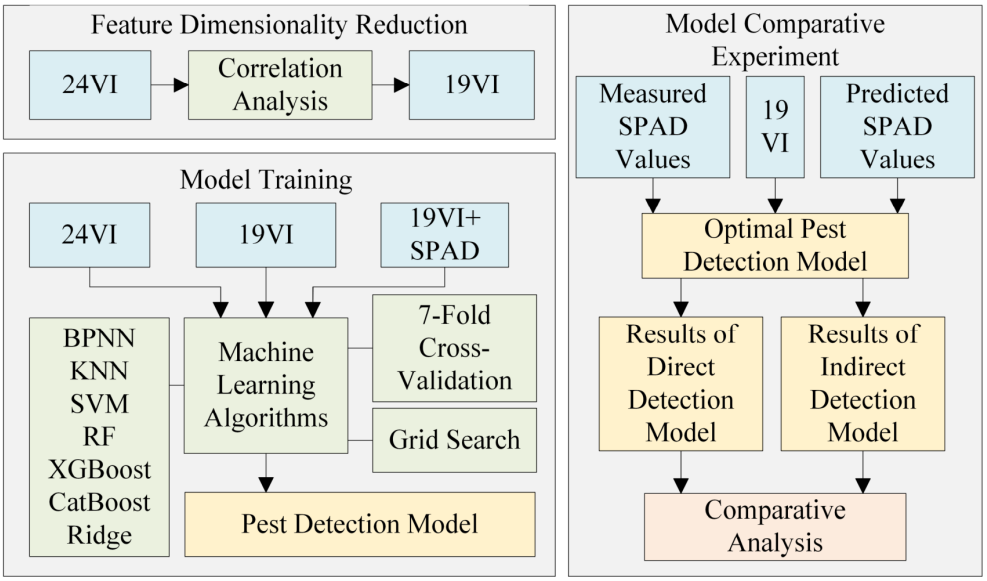


Figure 5. Technical workflow for constructing and validating the pest detection model.

As demonstrated by the comprehensive comparison in Table 5, the CatBoost-opt model utilizing the reduced feature set yielded the highest predictive accuracy. By excluding five vegetation indices showing relatively weak correlations with SPAD and applying Grid Search optimization, the model achieved an R^2 of 0.81 and an RMSE of 1.84. These results indicate that the reduced feature set maintained comparable predictive performance while decreasing the number of input variables. Internal model performance was further examined via 7-fold cross-validation (Figure 6), wherein the regression lines (solid red) across all folds closely tracked the 1:1 reference line (dashed). Exhibiting R^2 values consistently above 0.73 and exceeding 0.78 in the majority of folds, the CatBoost-opt model showed relatively stable internal performance across the seven validation folds. Consequently, the CatBoost-19opt model was selected for subsequent SPAD estimation analysis.

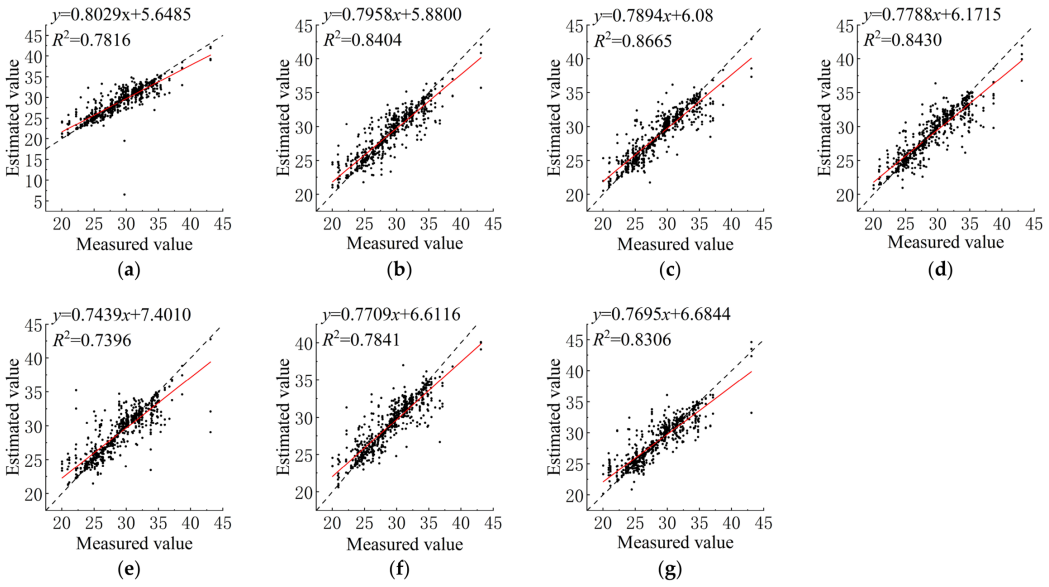


Figure 6. Cross-validation results: (**a**) Fold 1; (**b**) Fold 2; (**c**) Fold 3; (**d**) Fold 4; (**e**) Fold 5; (**f**) Fold 6; (**g**) Fold 7.

Across the seven validation folds, the CatBoost-19opt model achieved a mean R^2 of 0.81 ± 0.04 and an RMSE of 1.84 ± 0.11, indicating moderate fold-to-fold variability in internal model performance.

The independent test-set performance of all SPAD regression models is summarized in Table 8. Among the evaluated models, CatBoost achieved the highest independent test-set performance, with an R^2 of 0.84 and an RMSE of 1.56. Notably, the CatBoost-19opt model had already been selected during the model-development stage, and the independent test set was not used for feature screening, hyperparameter optimization, or model selection. Therefore, these results provide an independent evaluation of the predictive performance of the final SPAD regression model.

Table 8. Independent test-set performance of SPAD regression models.

Model	R^2	RMSE
BPNN	0.25	3.42
KNN	0.74	2.01
SVM	0.38	3.12
RF	0.71	2.11
XGBoost	0.67	2.27
CatBoost	0.84	1.56
Ridge	0.25	3.43

3.3.2. Comparative Analysis of *Eulecanium giganteum* Monitoring Models

Hyperparameter optimization is an important component of model development. As illustrated in Figure 7, hyperparameter optimization improved the internal classification performance of several models, with varying degrees of improvement. Notably, the Support Vector Machine and Extreme Gradient Boosting models exhibited high sensitivity to hyperparameter tuning. Specifically, the accuracy of SVM-19sopt increased from 0.63 to 0.88 (a 39% improvement over the baseline), while XGBoost-19sopt rose from 0.77 to 0.89 (a 15% improvement). These results indicate that hyperparameter optimization was associated with improved internal predictive performance within the parameter ranges evaluated in this study. Although Random Forest and CatBoost showed less pronounced post-tuning improvements, they demonstrated high baseline accuracy under default parameters (e.g., RF-24s achieved an accuracy of 0.86). Within the evaluated parameter ranges, these ensemble models showed relatively high accuracy under default settings, with comparatively small gains after tuning. In contrast, the linear nature of the Ridge model prevented it from effectively mapping the complex, non-linear boundaries between the high-dimensional spectral and physiological features. Consequently, it failed to achieve high accuracy even after extensive parameter tuning.

As illustrated in Figure 8, a comprehensive comparison of the seven machine learning models shows that model performance remained generally stable after reducing the feature set from 24 to 19 vegetation indices based on SPAD-correlation screening. Overall, the accuracy of the compared models changed little following feature reduction. For several models—specifically KNN, SVM, and CatBoost-opt—the Overall Accuracy remained completely unchanged even after five spectral variables were removed. Accuracy variations for most other models, including BPNN-24, CatBoost-24, and XGBoost-24s, were minimal (ranging from 0.01 to 0.02). This stability indicates that excluding vegetation indices with relatively weak correlations with SPAD did not markedly reduce model performance under the current modelling framework. Consequently, the reduced feature set decreased the number of input variables while maintaining comparable classification performance. Notably, within the single spectral feature category, SVM-19opt achieved a slightly higher OA (0.79) than SVM-24opt (OA = 0.78) (Figure 8b). Similarly, SVM-19sopt achieved an

OA of 0.88 compared with 0.87 for SVM-24sopt (Figure 8d). These results suggest that the reduced feature set can maintain comparable classification performance while decreasing the number of input variables.

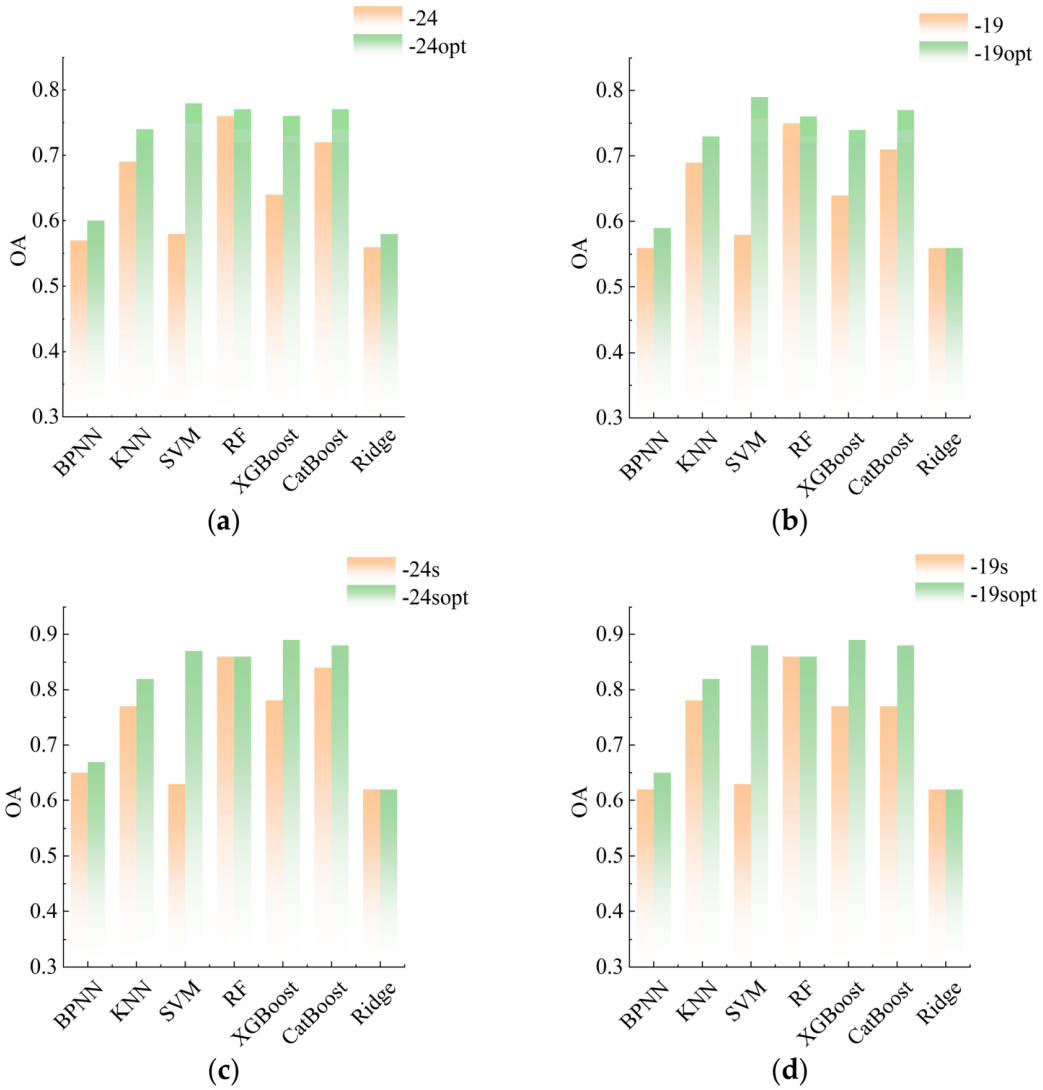


Figure 7. Comparison of parameter optimization results: model performance using (**a**) 24 VIs, (**b**) 19 VIs, (**c**) 24 VIs + SPAD, and (**d**) 19 VIs + SPAD as features.

Figure 9 illustrates the performance differences between models with and without multi-source feature fusion. Integrating SPAD as an additional feature improved the predictive accuracy of both the full-feature and reduced-feature models. Notably, the accuracy of the CatBoost-24 model rose from 0.72 to 0.84, representing a 16% increase following feature fusion. As shown in Table 6, models relying solely on spectral features exhibited limited performance even after optimization, with the best-performing SVM-19opt achieving an OA of 0.79 and a Kappa of 0.68. Under the current model-development framework, the multi-source feature configuration showed higher internal classification performance than the single-spectral feature models. Specifically, the XGBoost-19sopt model, which incorporates measured SPAD, achieved an OA of 0.89 and a Kappa coefficient of 0.84, whereas the XGBoost-19opt model achieved an OA of 0.74 and a Kappa coefficient of 0.62. These results indicate that incorporating measured SPAD information was associated with improved classification performance in the present dataset. Multispectral indices characterize canopy spectral responses, whereas SPAD provides physiological information related to chloro-

phyll status. Therefore, combining these features may provide complementary information for infestation classification under the current study conditions.

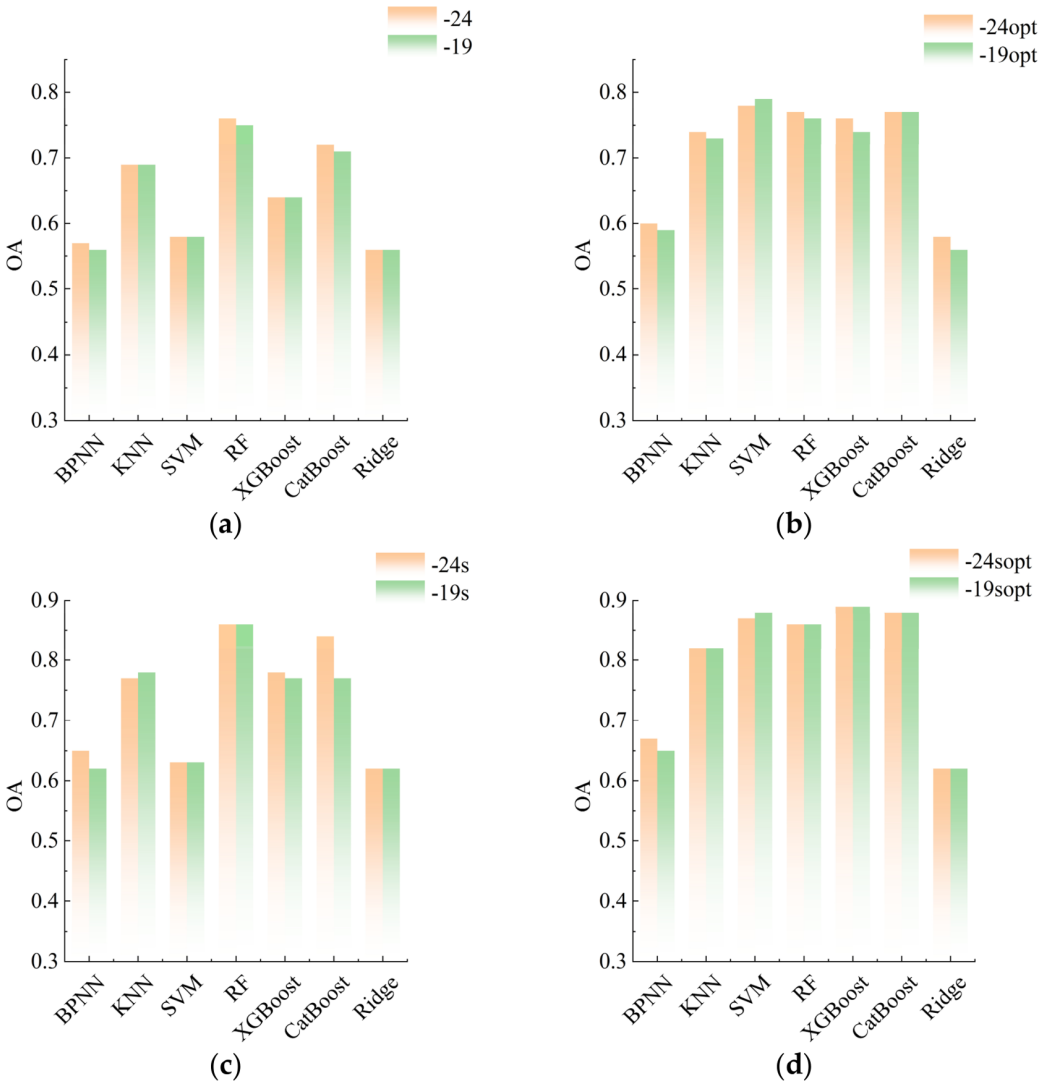


Figure 8. Comparison of model performance after feature dimensionality reduction: (**a**) original models using 24 VIs as features; (**b**) parameter-optimized models using 24 VIs as features; (**c**) original models using 24 VIs and SPAD as features; and (**d**) parameter-optimized models using 24 VIs and SPAD as features.

To evaluate the feasibility of the proposed framework, the best-performing XGBoost-based configurations were compared on the independent test set using three input configurations: 19VI alone, 19VI combined with measured SPAD, and 19VI combined with predicted SPAD. As detailed in Table 9, when utilizing measured SPAD values, the model achieved an OA of 0.88, a Kappa coefficient of 0.82, and an F1 of 0.88. In comparison, the VI-only model achieved an OA of 0.71, a Kappa coefficient of 0.55, and an F1 of 0.69. When predicted SPAD was incorporated, the OA remained at 0.71, while the Kappa coefficient increased slightly from 0.55 to 0.56 and the F1 increased from 0.69 to 0.71. Therefore, predicted SPAD did not improve OA relative to the VI-only model, although slight increases in Kappa and F1 were observed.

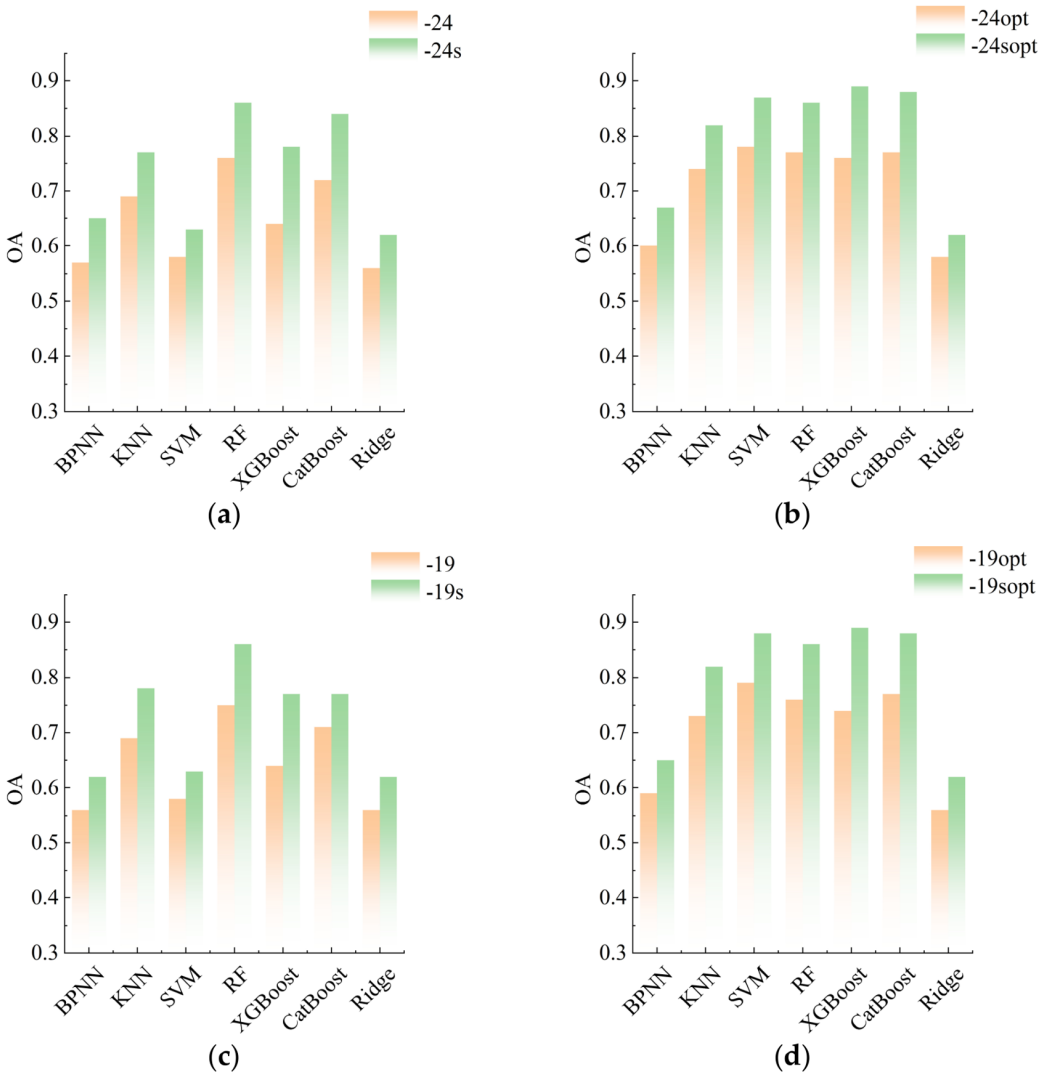


Figure 9. Comparison of multi-source fused features. Performance comparison between VI-only and fused VI + SPAD features across different sets: (**a**) 24 VIs; (**b**) 24 VIs with parameter optimization (24VI-opt); (**c**) 19 VIs; and (**d**) 19 VIs with parameter optimization (19VI-opt).

Table 9. Classification performance on the independent test set.

Model	Input Features	OA	Kappa	F1
XGBoost-19sopt	19VI + Measured SPAD	0.88	0.82	0.88
XGBoost-19sopt	19VI + Predicted SPAD	0.71	0.56	0.71
XGBoost-19opt	19VI	0.71	0.55	0.69

As shown in Table 10, the measured-SPAD configuration exhibited relatively balanced class-specific performance across the three infestation levels, with F1-scores of 0.91, 0.84, and 0.89 for the Healthy, Mild, and Severe classes, respectively. When predicted SPAD was used, class-specific performance decreased, with F1-scores of 0.74, 0.65, and 0.74, respectively. The Mild class showed the lowest F1-score, indicating greater difficulty in distinguishing the intermediate infestation level.

Table 10. Class-specific classification performance of the XGBoost-19sopt model on the independent test set.

Input Features	Class	Precision	Recall	F1
19VI + Measured SPAD	Healthy	0.94	0.89	0.91
	Mild	0.82	0.86	0.84
	Severe	0.89	0.89	0.89
19VI + Predicted SPAD	Healthy	0.76	0.72	0.74
	Mild	0.64	0.67	0.65
	Severe	0.74	0.74	0.74

To further evaluate the classification performance of the model across different infestation levels, the classification results on the independent test set were analyzed using a normalized confusion matrix (Figure 10). The results showed that, when predicted SPAD values were used, the Healthy and Severe classes showed relatively higher recognition rates, whereas the Mild class exhibited greater confusion with the adjacent infestation levels. Misclassifications occurred primarily between adjacent infestation levels, indicating greater difficulty in distinguishing the intermediate Mild category.

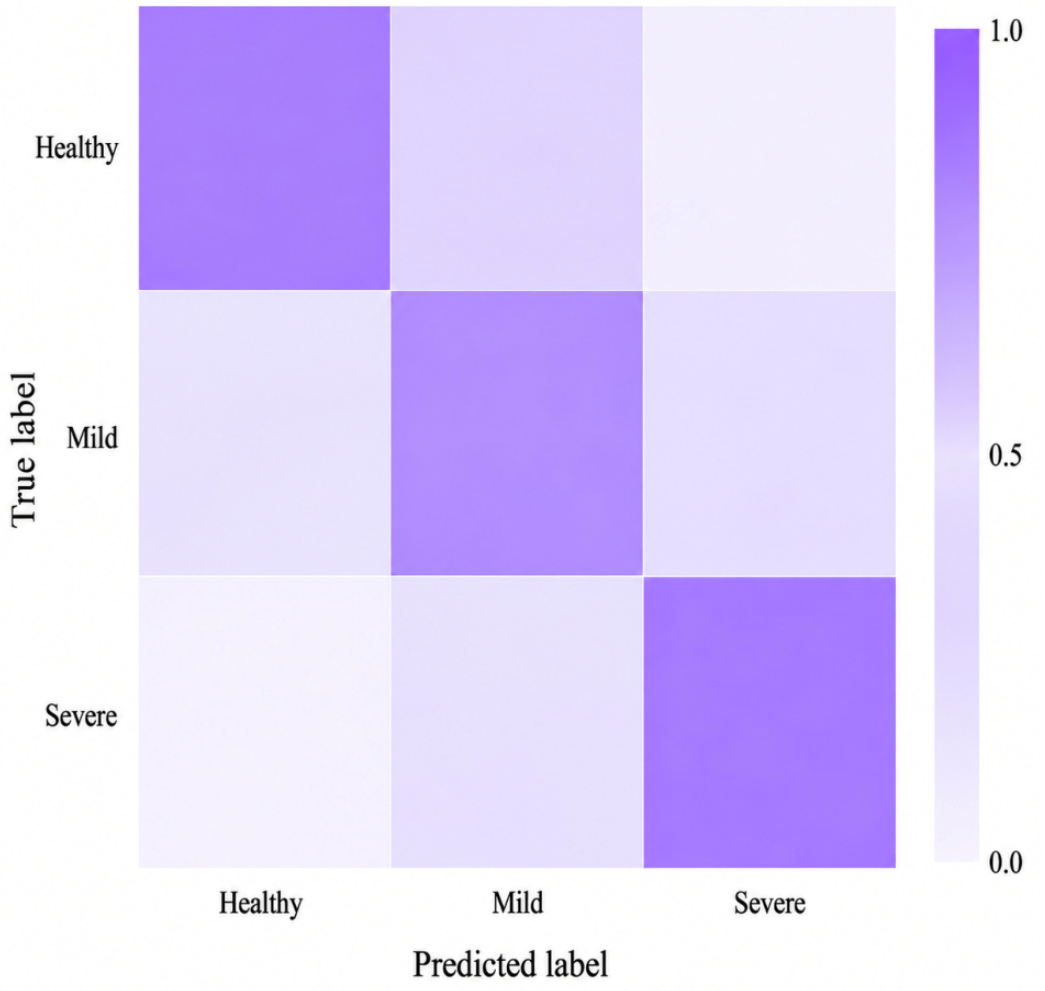


Figure 10. Normalized confusion matrix of the XGBoost-19sopt model for Healthy, Mild, and Severe infestation classification using predicted SPAD values on the independent test set.

4. Discussion

4.1. Effect of Feature Reduction and Parameter Tuning on Model Performance

When working with high-dimensional and multi-source features, feature reduction and parameter tuning are crucial for developing stable and efficient models. In this study, five vegetation indices showing relatively weak correlations with SPAD ($|r| < 0.30$) were excluded through Pearson correlation screening. The results demonstrate that models built on the reduced feature set maintained stable accuracy while reducing the number of input variables. Notably, the predictive accuracy of certain models, such as SVM-19sopt, actually improved. This suggests that excluding vegetation indices with relatively weak correlations with SPAD did not markedly reduce model performance under the current modelling framework. However, because the screening criterion was based on the correlation between each vegetation index and SPAD rather than on inter-feature redundancy, the excluded variables should not be interpreted as redundant or highly collinear noise. Therefore, Pearson correlation screening provided a more compact feature set for subsequent model development. Furthermore, hyperparameter optimization played an important role in model development. Tree-based models (CatBoost and XGBoost) and kernel-based models exhibited varying degrees of accuracy improvement following Grid Search optimization. These results indicate that the performance of these models was sensitive to hyperparameter configurations. Proper parameter optimization improved their internal model performance within the evaluated feature space. However, because feature screening and hyperparameter selection were conducted within the overall model-development workflow without nested grouped cross-validation, these results should be interpreted as internal model-development performance rather than fully independent estimates of model generalization.

4.2. Advantages of Multi-Source Feature Fusion in Pest Monitoring Model Construction

Eulecanium giganteum feeds by piercing and sucking sap from its host plants. This prolonged biotic stress can adversely affect chlorophyll synthesis and accelerate chlorophyll degradation in jujube leaves. Canopy spectral features can reflect changes in canopy structure and overall growth. However, their ability to identify mild stress may be limited under the current study conditions. By comparison, SPAD values serve as a proxy for relative chlorophyll content and provide complementary information about leaf chlorophyll status. By incorporating SPAD values to construct models with multi-source features, the recognition accuracy of all classification models improved. Most notably, the Overall Accuracy of the optimal model, XGBoost-19sopt, increased from 0.74 to 0.89. This finding aligns with previous research on the remote sensing monitoring of agricultural pests and diseases [54,55]. Therefore, combining measured SPAD values with vegetation indices provided complementary spectral and physiological information and was associated with higher observed classification performance in the present dataset.

It should be noted that SPAD values and canopy spectral characteristics may also be influenced by factors such as water and nutrient availability, phenological stage, and canopy structure. Therefore, the spectral and physiological responses observed in this study should be interpreted as being associated with *Eulecanium giganteum* infestation rather than as strictly pest-specific signatures. Future studies incorporating additional environmental variables and multi-site validation will help further improve the robustness and generalizability of the proposed monitoring framework.

4.3. Analysis of Experimental Results for Indirect Monitoring

While in situ measured SPAD values provide highly accurate baseline data, their reliance on manual sampling makes them impractical for efficient, large-scale orchard mon-

itoring. To address this limitation, this study evaluated an indirect monitoring approach in which measured SPAD values were replaced with values predicted by the regression model. It should be emphasized that Predicted SPAD is derived from the same vegetation indices subsequently used in the classification model and therefore does not constitute an independent information source. Instead, it should be regarded as a model-derived physiological representation of the multispectral predictors within a two-stage indirect monitoring framework. Although the test results indicated a decline in classification performance, the indirect monitoring model achieved an Overall Accuracy (OA) of 0.71. This decrease in classification performance may be associated with the uncertainty introduced by Predicted SPAD values. When these estimated features were introduced into the classification model, prediction errors could propagate through the feature space and affect the resulting decision boundaries, thereby reducing classification accuracy. Compared with the VI-only model, the incorporation of Predicted SPAD did not increase OA, which remained at 0.71. The Kappa coefficient increased slightly from 0.55 to 0.56, whereas the F1 increased from 0.69 to 0.71. These results indicate that Predicted SPAD did not provide a clear improvement in overall classification performance on the independent test set, but it retained potential as a physiologically interpretable intermediate representation within the two-stage indirect monitoring framework. These results support the proposed approach as a proof of concept under the current study conditions, while further model optimization and external validation across different orchards, years, and phenological stages are required before large-scale operational application.

5. Conclusions

In this study, UAV multispectral vegetation indices and SPAD values were integrated with machine learning algorithms to assess *Eulecanium giganteum* infestation in jujube orchards. The main conclusions are as follows:

(1) Pearson correlation-based screening reduced the original 24 vegetation indices to a subset of 19 indices showing stronger associations with SPAD. During internal model development, the CatBoost-19opt model achieved an R^2 of 0.81 and an RMSE of 1.84. On the independent test set of 58 jujube trees, the final model achieved an R^2 of 0.84 and an RMSE of 1.56, providing an independent assessment of its predictive performance for tree-level SPAD estimation.

(2) The incorporation of measured SPAD values improved pest classification performance compared with the use of vegetation indices alone. Among the evaluated models, XGBoost-19sopt achieved the best overall performance, with an OA of 0.89 and a Kappa coefficient of 0.84 during model comparison. On the independent test set, the model using measured SPAD values achieved an OA of 0.88, a Kappa coefficient of 0.82, and a F1 of 0.88.

(3) When measured SPAD values were replaced with Predicted SPAD values, classification performance on the independent test set decreased to an OA of 0.71, a Kappa coefficient of 0.56, and an F1 of 0.71. Compared with the VI-only model (OA = 0.71, Kappa = 0.55, and F1 = 0.69), the addition of Predicted SPAD did not improve OA, although slight increases were observed in Kappa and F1. These findings indicate that Predicted SPAD does not currently provide a clear improvement in overall classification performance, but it may serve as a physiologically interpretable intermediate representation within the two-stage indirect monitoring framework.

Author Contributions: Conceptualization, Y.C., R.H., R.Y., H.D. and P.L.; Methodology, Y.C., Z.X., R.Y. and P.L.; Software, Y.C.; Validation, Y.C. and Y.W.; Formal analysis, Y.C., Z.X., R.H. and H.D.; Investigation, Y.C., Y.W., Z.X., R.H., R.Y. and H.D.; Resources, P.L.; Data curation, Y.C., Y.W., Z.X. and

H.D.; Writing—original draft, Y.C. and Y.W.; Writing—review & editing, Y.C. and P.L.; Visualization, Y.C. and Y.W.; Supervision, P.L.; Project administration, P.L.; Funding acquisition, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Corps "Three Rural Issues" Backbone Talent Training Project (No. 2025SNGGGCC-BT09), entitled "Research and Development of Key Technologies and Equipment for Intelligent Plant Protection Machinery in Xinjiang Jujube Orchards".

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bai, T.; Zhang, N.; Mercatoris, B.; Chen, Y. Jujube yield prediction method combining Landsat 8 Vegetation Index and the phenological length. *Comput. Electron. Agric.* **2019**, *162*, 1011–1027. [CrossRef]
2. Zhu, X.; Xu, B.; Kadeer, A.; Li, H.; Song, B.; Chen, H.; Hao, J. The occurrence and control status of fruit tree diseases and insect pests in Xinjiang: A case study of jujube. *Xinjiang Agric. Sci.* **2024**, *61*, 116–121. (In Chinese) [CrossRef]
3. Makan, N.; Kong, T.; Ayixiamu; Liu, A. Occurrence, damage and control measures of *Eulecanium giganteum* in Xinjiang forest areas. *South China Agric.* **2020**, *14*, 31–32. (In Chinese) [CrossRef]
4. Zhang, S.; Wang, X.; Lin, H.; Dong, Y.; Qiang, Z. A review of the application of UAV multispectral remote sensing technology in precision agriculture. *Smart Agric. Technol.* **2025**, *12*, 101406. [CrossRef]
5. Glenn, E.P.; Neale, C.M.U.; Hunsaker, D.J.; Nagler, P.L. Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems. *Hydrol. Process.* **2011**, *25*, 4050–4062. [CrossRef]
6. Chen, T.; Zeng, R.; Guo, W.; Hou, X.; Lan, Y.; Zhang, L. Detection of Stress in Cotton (*Gossypium hirsutum* L.) Caused by Aphids Using Leaf Level Hyperspectral Measurements. *Sensors* **2018**, *18*, 2798. [CrossRef]
7. Gonçalves, F.V.J.; Falcioni, R.; Rutz, T.; da Silva, A.L.B.R.; Furlanetto, R.H.; Crusiol, L.G.T.; de Oliveira, K.M.; de Oliveira, C.A.; Vedana, N.G.; Demattê, J.A.M.; et al. Mapping of Leaf Pigments in Lettuce via Hyperspectral Imaging and Machine Learning. *Horticulturae* **2025**, *11*, 1077. [CrossRef]
8. Yao, Y.; Wang, D.; Shu, M.; Hu, J.; Liu, Y.; Feng, H.; Li, B.; Lu, W.; Wang, J.; Xi, L.; et al. Estimation of maize LCC and FVC for full growth periods based on UAV remote sensing and ensemble learning. *PeerJ Comput. Sci.* **2025**, *11*, e3356. [CrossRef]
9. Zhuang, W.; Li, D.; Kou, W.; Lu, N.; Wu, F.; Sun, S.; Liu, Z. A Multi-Feature Estimation Model for Olive Canopy Chlorophyll Combining XGBoost with UAV Imagery. *Agronomy* **2025**, *15*, 2718. [CrossRef]
10. Huang, Y.; Li, D.; Liu, X.; Ren, Z. Monitoring canopy SPAD based on UAV and multispectral imaging over fruit tree growth stages and species. *Front. Plant Sci.* **2024**, *15*, 1435613. [CrossRef]
11. Zhang, S.; Huang, X.; Ma, L.; Dashzebeg, G.; Ariunaa, M.; Bao, G.; Tong, S.; Bao, Y.; Dorjsuren, A.; Enkhnasan, D. Response of spectral vegetation indices to *Erannis jacobsoni* Djak. damage in larch forests. *Ecol. Inform.* **2024**, *81*, 102605. [CrossRef]
12. Yu, H.; Hu, L.; Cui, W.; Yang, L.; Li, J.; Han, G.; Zhou, Q.; Wang, Z.; Lan, Y.; Zhao, J.; et al. A novel vegetation index for monitoring the stress levels of pest caused by dusky cotton bug. *Comput. Electron. Agric.* **2025**, *234*, 110221. [CrossRef]
13. Chen, Z.; Zhai, W.; Cheng, Q. Enhancing maize LAI estimation accuracy using unmanned aerial vehicle remote sensing and deep learning techniques. *Artif. Intell. Agric.* **2025**, *15*, 482–495. [CrossRef]
14. Liu, Z.; Ding, D.; Hamani, A.K.M.; Zhai, W.; Tian, J.; Wang, Y.; Du, Y. Estimation of soybean phenotypic parameters across growth stages using UAV-based multi-source feature fusion and XGBoost. *Clim. Smart Agric.* **2026**, *3*, 100098. [CrossRef]
15. Ai, W.; Yang, G.; Li, Z.; Du, J.; Ye, L.; Feng, X.; He, Y. Two-year remote sensing and ground verification: Estimating chlorophyll content in winter wheat using UAV multi-spectral imagery. *Artif. Intell. Agric.* **2026**, *16*, 480–494. [CrossRef]
16. Wang, S.; Wang, B.; Zheng, J.; Tuerxun, N.; Wang, R.; Zhang, K.; Yang, Y. Contributions of Multiple UAV Features to Cotton SPAD Estimation from the Perspective of Explainable Machine Learning. *Agriculture* **2026**, *16*, 325. [CrossRef]
17. Yan, N.; Qin, Y.; Wang, H.; Wang, Q.; Hu, F.; Wu, Y.; Zhang, X.; Li, X. The Inversion of SPAD Value in Pear Tree Leaves by Integrating Unmanned Aerial Vehicle Spectral Information and Textural Features. *Sensors* **2025**, *25*, 618. [CrossRef]
18. He-Ya, S.; Huang, X.; Zhou, D.; Zhang, J.; Bao, G.; Tong, S.; Bao, Y.; Ganbat, D.; Tsagaantsooj, N.; Altanchimeg, D.; et al. Identification of Larch Caterpillar Infestation Severity Based on Unmanned Aerial Vehicle Multispectral and LiDAR Features. *Forests* **2024**, *15*, 191. [CrossRef]
19. Hornero, A.; Zarco Tejada, P.J.; Marengo, I.; Faria, N.; Hernández Clemente, R. Detection of oak decline using radiative transfer modelling and machine learning from multispectral and thermal RPAS imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *127*, 103679. [CrossRef]

20. Wu, J.; Bai, T.; Li, X. Inverting Chlorophyll Content in Jujube Leaves Using a Back-Propagation Neural Network–Random Forest–Ridge Regression Algorithm with Combined Hyperspectral Data and Image Color Channels. *Agronomy* **2024**, *14*, 140. [CrossRef]
21. Tuerxun, N.; Zheng, J.; Wang, R.; Wang, L.; Liu, L. Hyperspectral estimation of chlorophyll content in jujube leaves: Integration of derivative processing techniques and dimensionality reduction algorithms. *Front. Plant Sci.* **2023**, *14*, 1260772. [CrossRef]
22. Wang, L.; Zeng, J.; Yu, M.; Fan, W.; Bao, J. Detection of SPAD Content in Leaves of Grey Jujube Based on Near Infrared Spectroscopy. *Horticulturae* **2025**, *11*, 1251. [CrossRef]
23. Liu, G.; He, J.; Wang, S.; Luo, Y.; Wang, W.; Wu, L.; Si, Z.; He, X. Application of Near-Infrared Hyperspectral Imaging for Detection of External Insect Infestations on Jujube Fruit. *Int. J. Food Prop.* **2016**, *19*, 41–52. [CrossRef]
24. Jiang, M.; Li, Y.; Song, J.; Wang, Z.; Zhang, L.; Song, L.; Bai, B.; Tu, K.; Lan, W.; Pan, L. Study on Black Spot Disease Detection and Pathogenic Process Visualization on Winter Jujubes Using Hyperspectral Imaging System. *Foods* **2023**, *12*, 435. [CrossRef]
25. Bulacio Fischer, P.T.; Carella, A.; Massenti, R.; Fadhilah, R.; Lo Bianco, R. Advances in Monitoring Crop and Soil Nutrient Status: Proximal and Remote Sensing Techniques. *Horticulturae* **2025**, *11*, 182. [CrossRef]
26. *LY/T 1681-2006*; Standard of Forest Pests Occurrence and Disaster. Standards Press of China: Beijing, China, 2006.
27. Zhu, H.; Huang, Y.; An, Z.; Zhang, H.; Han, Y.; Zhao, Z.; Li, F.; Zhang, C.; Hou, C. Assessing radiometric calibration methods for multispectral UAV imagery and the influence of illumination, flight altitude and flight time on reflectance, vegetation index and inversion of winter wheat AGB and LAI. *Comput. Electron. Agric.* **2024**, *220*, 108821. [CrossRef]
28. Lu, C.; Dong, S.; Yu, X.; Li, Y. A UAV-Based Salt Stress Response Index for Comparative Evaluation of Salinity-Management Interventions in Maize and Soybean. *Agriculture* **2026**, *16*, 1769. [CrossRef]
29. Gao, P.; Wang, K.; Li, Y.; Zhang, J.; Li, Q.; Lu, W.; Song, Y. An Artificial Intelligence-Driven UAV and Ground Sensor Fusion Framework for Crop Growth Assessment in Smart Agriculture. *Agriculture* **2026**, *16*, 1650. [CrossRef]
30. Zhou, Z.; Cui, D.; Fu, J.; Li, M.; Zhu, S.; Zhang, L.; Yu, R.; Qiu, Q.; Chen, X.; Chen, Q.; et al. Multi-Temporal Prediction of Soybean Yield at the Plot Scale Through Multi-Source UAV Sensor Fusion and Dynamic Modeling. *Agriculture* **2026**, *16*, 1548. [CrossRef]
31. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [CrossRef]
32. Ye, Y.; Wu, M.-T.; Pu, C.-J.; Chen, J.-M.; Jing, Z.-X.; Shi, T.-T.; Zhang, X.-B.; Yan, H. Monitoring Chrysanthemum Cultivation Areas Using Remote Sensing Technology. *Horticulturae* **2025**, *11*, 933. [CrossRef]
33. Niu, X.; Chen, B.; Sun, W.; Feng, T.; Yang, X.; Liu, Y.; Liu, W.; Fu, B. Estimation of Coastal Wetland Vegetation Aboveground Biomass by Integrating UAV and Satellite Remote Sensing Data. *Remote Sens.* **2024**, *16*, 2760. [CrossRef]
34. Yang, P.; van der Tol, C. Linking canopy scattering of far-red sun-induced chlorophyll fluorescence with reflectance. *Remote Sens. Environ.* **2018**, *209*, 456–467. [CrossRef]
35. Kahriman, A.; Günlü, A.; Karahalil, U. Estimation of Crown Closure and Tree Density Using Landsat TM Satellite Images in Mixed Forest Stands. *J. Indian Soc. Remote Sens.* **2014**, *42*, 559–567. [CrossRef]
36. Liu, H.; Zhang, F.; Zhang, L.; Lin, Y.; Wang, S.; Xie, Y. UNVI-Based Time Series for Vegetation Discrimination Using Separability Analysis and Random Forest Classification. *Remote Sens.* **2020**, *12*, 529. [CrossRef]
37. Cheng, J.; Yang, H.; Qi, J.; Sun, Z.; Han, S.; Feng, H.; Jiang, J.; Xu, W.; Li, Z.; Yang, G. Estimating canopy-scale chlorophyll content in apple orchards using a 3D radiative transfer model and UAV multispectral imagery. *Comput. Electron. Agric.* **2022**, *202*, 107401. [CrossRef]
38. Xing, N.; Huang, W.; Xie, Q.; Shi, Y.; Ye, H.; Dong, Y.; Wu, M.; Sun, G.; Jiao, Q. A Transformed Triangular Vegetation Index for Estimating Winter Wheat Leaf Area Index. *Remote Sens.* **2019**, *12*, 16. [CrossRef]
39. Hamzaoui, H.; Maafa, I.; Choukri, H.; Bakkali, A.E.; Houssaini, S.E.I.E.; Razouk, R.; Aziz, A.; Louahlia, S.; Habbadi, K. Assessment of Peacock Spot Disease (*Fusicladium oleagineum*) in Olive Orchards Through Agronomic Approaches and UAV-Based Multispectral Imaging. *Horticulturae* **2025**, *11*, 46. [CrossRef]
40. Choosumrong, S.; Hataitara, R.; Sujipuli, K.; Weerawatanakorn, M.; Preechaharn, A.; Premjet, D.; Laywisadkul, S.; Raghavan, V. Bananas diseases and insect infestations monitoring using multi-spectral camera RTK UAV images. *Spat. Inf. Res.* **2023**, *31*, 243–255. [CrossRef]
41. Egea-Cobrero, V.; Rodriguez-Galiano, V.; Sánchez-Rodríguez, E.; García-Pérez, M.A. Estimación de la cosecha de trigo en Andalucía usando series temporales de MERIS Terrestrial Chlorophyll Index (MTCI). *Rev. Teledetec.* **2018**, *51*, 11–24. [CrossRef]
42. Romano, E.; Bergonzoli, S.; Pecorella, I.; Bisaglia, C.; De Vita, P. Methodology for the Definition of Durum Wheat Yield Homogeneous Zones by Using Satellite Spectral Indices. *Remote Sens.* **2021**, *13*, 2036. [CrossRef]
43. Guermazi, E.; Wali, A.; Ksibi, M. Combining remote sensing, SPAD readings, and laboratory analysis for monitoring olive groves and olive oil quality. *Precis. Agric.* **2024**, *25*, 65–82. [CrossRef]
44. Chen, C.; Yuan, X.; Gan, S.; Luo, W.; Bi, R.; Li, R.; Gao, S. A new vegetation index based on UAV for extracting plateau vegetation information. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *128*, 103668. [CrossRef]

45. Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 235–248. [CrossRef]
46. Ekwe, M.C.; Adeluyi, O.; Verrelst, J.; Umar, I.Y.; Omogunloye, J.O.; Shaba, H.A. Estimating rainfed groundnut's leaf area index using Sentinel-2 based on Machine Learning Regression Algorithms and Empirical Models. *Precis. Agric.* **2024**, *25*, 1404–1428. [CrossRef]
47. Mohammadi, S.; Uhlen, A.K.; Lillemo, M.; Giezendanner, R.; Ergon, Å. Enhancing phenotyping efficiency in faba bean breeding: Integrating UAV imaging and machine learning. *Precis. Agric.* **2024**, *25*, 1502–1528. [CrossRef]
48. Wang, X.; Wu, S.; Han, L.; Wang, J.; Cui, J.; Shen, Y. Slope stability prediction based on GSOEM-SV: A mobile application practicably deploy in engineering verification. *Adv. Eng. Softw.* **2024**, *189*, 103648. [CrossRef]
49. Jiao, Z.; Zhang, D.; Zhang, J.; Wang, L.; Ma, D.; Ma, L.; Wang, Y.; Gu, A.; Fan, X.; Peng, B.; et al. Early Detection of Chinese Cabbage Clubroot Based on Integrated Leaf Multispectral Imaging and Machine Learning. *Horticulturae* **2025**, *11*, 1335. [CrossRef]
50. Zhai, W.; Li, C.; Fei, S.; Liu, Y.; Ding, F.; Cheng, Q.; Chen, Z. CatBoost algorithm for estimating maize above-ground biomass using unmanned aerial vehicle-based multi-source sensor data and SPAD values. *Comput. Electron. Agric.* **2023**, *214*, 108306. [CrossRef]
51. Awichin, P.; Laosuwan, T.; Sangpradid, S.; Uttaruk, Y.; Butthep, C.; Intarat, K.; Laoratthaphong, N.; Phoophathong, T.; Jeefoo, P.; Singharaj, M. Hybrid Feature Selection and Ensemble Learning for Aboveground Carbon Mapping in Oil Palm Plantations Using Multi-Source Satellite Data. *Agriculture* **2026**, *16*, 1834. [CrossRef]
52. Wang, X.; Yang, W.; Yang, Y.; Huang, M.; Zhu, Q. Identification of tomato bacterial wilt severity based on hyperspectral imaging technology and spectrum Transformer network. *Ecol. Inform.* **2023**, *78*, 102353. [CrossRef]
53. Cocozza, A.; Venezia, A.; Macellaro, R.; Di Cesare, C.; Milanesi, C.; Tripodi, P. Leveraging Multispectral and 3D Phenotyping to Determine Morpho-Physiological Changes in Peppers Under Increasing Drought Stress Levels. *Horticulturae* **2025**, *11*, 1318. [CrossRef]
54. Grishina, A.; Lysov, M.; Ageyeva, M.; Diakova, V.; Sherstneva, O.; Brilkina, A.; Vodeneev, V. Bacterial and Viral-Induced Changes in the Reflectance Spectra of *Nicotiana benthamiana* Plants. *Horticulturae* **2024**, *10*, 1363. [CrossRef]
55. Zheng, Q.; Huang, W.; Xia, Q.; Dong, Y.; Ye, H.; Jiang, H.; Chen, S.; Huang, S. Remote Sensing Monitoring of Rice Diseases and Pests from Different Data Sources: A Review. *Agronomy* **2023**, *13*, 1851. [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.