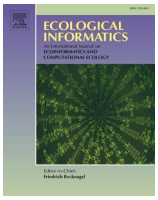
Ecological Informatics 90 (2025) 103301

Contents lists available at ScienceDirect

Ecological Informatics

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



Assessing *Gonipterus* defoliation levels using multispectral unmanned aerial vehicle (UAV) data in *Eucalyptus* plantations

Phumlani Nzuza [a,b,*], Michelle L. Schröder [a], Rene J. Heim [b,c], Louis Daniels [b], Bernard Slippers [d], Brett P. Hurley [a], IIaria Germishuizen [e], Benice Sivparsad [e], Jolanda Roux [f,g], Wouter. H Maes [b]

[a] Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa
[b] UAV research centre (URC), Department of Plants and Crops, Ghent University, Belgium
[c] Photogrammetry & Robotics Lab, University of Bonn, Nussallee 15, 53115 Bonn, Germany
[d] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa
[e] Institute for Commercial Forestry Research (ICFR), PO Box 100281, Scottsville, Pietermaritzburg, South Africa
[f] Sappi Forests, RPN, Shaw Research Centre, Howick, RSA, South Africa
[g] Department of Plant and Soil Science, Forestry and Agricultural Biotechnology of Institute (FABI), University of Pretoria, South Africa

ARTICLE INFO

Keywords:
Invasive insects
XGboost classifier
SVM
RF
Image texture
Forest entomology

ABSTRACT

Invasive insect pest *Gonipterus* sp. n. 2 (Coleoptera: Curculionidae) threatens *Eucalyptus* species, causing defoliation and yield loss through adult and larval feeding. Early detection is important for early intervention to prevent pest outbreaks. As conventional insect pest monitoring methods are time-consuming and spatially restrictive, this study assessed the potential of UAV monitoring. Multispectral imagery was obtained with Unmanned Aerial Vehicles (UAVs) across six different stands of young *Eucalyptus dunnii* with varying levels of *Gonipterus* sp. n. 2 infestations. Some stands were revisited, a total of 9 datasets were covered. Reference damage levels were obtained through visual assessments of (n = 89–100) trees at each site. Across sites, a decrease in canopy reflectance in both the visual and the near-infrared domains with increasing damage levels was consistently observed. Several vegetation indices showed consistent patterns, but none showed site independence. XGBoost, Support Vector Machine and Random Forest (RF) were used to predict damage levels using five input spectral data types. XGBoost performed best, closely followed by RF. Both models consistently selected very similar features. The best-performing models included reflectance, vegetation indices and grey-level co-occurrence matrix data. When data from 10 different wavelengths were used, the highest classification accuracy was 92 % across all sites in classifying defoliation levels. With a classical 5-band multispectral camera, accuracy was 88 %, but distinguishing medium damage from low remained challenging. However, the method was less reliable when trained and validated on separate fields. This study highlights the potential of multi-site datasets in increasing the model's generalization, using UAV based multispectral imagery to assess *Gonipterus* sp. n. 2 damage and demonstrating reliable upscaling from individual tree assessments to stand scale. However, it also recognises the difficulty of generating a robust model that performs well on untrained sites.

1. Introduction

Insect disturbances on forest health are associated with large economic losses for the forestry sector (Bradshaw et al., 2016; Canelles et al., 2021; Senf et al., 2017; Williams et al., 2023). Among the insect pests that pose significant threats to *Eucalyptus* (Myrtaceae) plantations, the *Eucalyptus* Snout Beetle (ESB; *Gonipterus* spp., Coleoptera: Curculionidae) has emerged as a species of concern. Native to south-east Australia and Tasmania, and exclusively feeding on *Eucalyptus* spp., the beetle has spread to numerous countries on five continents (Schröder et al., 2020). In South Africa, its initial detection dates back to 1916 (Tooke, 1955). ESB refers to three species; *Gonipterus platensis, G. pulverulentus* and an undescribed *Gonipterus* sp. n. 2, which all belong to the *G. scutellatus* species complex (Mapondera et al., 2012; Schröder et al., 2020). In South Africa, recent reports show that only *Gonipterus* sp. n. 2 is present (Mapondera et al., 2012). *E. dunnii*, one of the most widely

* Corresponding author at: Department of Zoology and Entomology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa.
E-mail address: phumlanipollen.nzuza@ugent.be (P. Nzuza).

https://doi.org/10.1016/j.ecoinf.2025.103301
Received 21 November 2024; Received in revised form 23 June 2025; Accepted 23 June 2025
Available online 24 June 2025

planted *Eucalyptus* species in South Africa (Morris, 2022), is susceptible to *Gonipterus* sp. n. 2 (Joubert et al., 2023; Tooke, 1955), and so are various other important economic *Eucalyptus* species, including *E. globulus*, *E. viminalis* and *E. smithii* (Mally, 1924; Newete et al., 2011; Tooke, 1955).

Feeding damage by *Gonipterus* spp. causes significant growth declines and yield losses (Tooke, 1955) with both larvae and adult stages of the snout weevil contributing to defoliation (Schröder et al., 2020). The larvae induce the most damage: the early instar larvae feed on the epidermis, whereas later instars consume entire young leaves (Tooke, 1955). Adult beetles primarily target the edges of mature leaves, leading to a distinctive scalloped appearance. Reis et al. (2012), estimated up to 86 % wood volume loss from *Gonipterus platensis* at 100 % defoliation levels, while Keetse (2024) estimated a 17.5 % wood volume loss from *Gonipterus* sp. n. 2 from just 18 % defoliation.

Within South Africa, efforts to suppress *Gonipterus* sp. n. 2 populations through classical biological control were initiated in 1926 with the successful release of *Anaphes nitens* Girault (Hymenoptera: Mymaridae), an egg parasitoid native to Australia (Tooke, 1955). By 1950, the population of *Gonipterus* sp. n. 2 was largely suppressed (Tooke, 1955). However, despite this initial success, sporadic outbreaks persist, even in the presence of the biological control agent (Tribe, 2005). The reasons for these sporadic outbreaks are not fully understood, but it is hypothesized that environmental factors and climate mismatch between the pest and parasitoid, and/or variation in host resistance, contribute to variable parasitism rates (Schröder et al., 2020). For instance, in colder or high-altitude areas in South Africa, *Gonipterus* sp. n. 2 activity decreases during winter, which on its turn reduces *Anaphes* (Tooke, 1955; Tribe, 2005). In spring, *Gonipterus* sp. n. 2 can build up quickly in numbers, before *Anaphes* does (Schröder et al., 2017; Tooke, 1955), resulting in sporadic disease outbreaks. Later in the season, high parasitism tends to avoid outbreaks. In low-altitude regions where winter temperatures are mild, *Gonipterus* sp. n. 2 remains active year-round, and so does *Anaphes*.

Monitoring tree health in forests and plantations is crucial for early detection and prevention, playing a vital role in biosecurity and effective pest management (Brockerhoff et al., 2023; Mngadi et al., 2024). Traditional methods, such as ground-based measurements and visual inspections, are often time-consuming, labour-intensive, and limited in spatial coverage. In contrast, remote sensing offers a powerful alternative for monitoring, detecting, and mapping vegetation damage and defoliation caused by insects (Debeurs and Townsend, 2008; Ismail and Mutanga, 2010). By capturing spectral changes linked to shifts in leaf composition or foliage loss (Lausch et al., 2016), remote sensing enables the early detection of pests and pathogens, allowing for timely intervention (Stone and Mohammed, 2017).

Some studies have confirmed the potential of satellite remote sensing for damage assessment of *Gonipterus* sp. n. 2. Lottering and Mutanga (2016) utilized vegetation indices derived from pan-sharpened Worldview 2 imagery to predict canopy damage induced by *Gonipterus* sp. n. 2. Expanding upon this, Lottering et al. (2019) further improved the detection of *Gonipterus* sp. n. 2. defoliation by incorporating image texture through grey level co-occurrence matrix analysis using artificial neural networks. Xulu et al. (2024) compared multiple spectral vegetation indices in detecting *Gonipterus* sp. n. 2 damage and other forest disturbances using anomaly detection and Landsat timeseries data. Their results demonstrated that the Normalized Difference Infra-red (NDII) and Enhanced Vegetation Index (EVI) could detect anomalies linked to *Gonipterus* sp. n. 2 damage. Likewise, Adelabu et al. (2014) explored the potential of the red-edge band from RapidEye to assess insect-induced damage levels within an African savanna using both RF and Support Vector Machine (SVM) classifiers, where SVM performed best, achieving 81 % overall classification accuracy. Dos Santos et al. (2017), using Landsat 8 data to investigate damage by *T. peregrinus* of *Eucalyptus urophylla x grandis* hybrid clones, found that the spectral reflectance in the visual to near-infrared spectrum was lower for healthy trees than for damaged trees. Together, these studies emphasize the capacity of satellite data to map insect damage. However, inherent limitations of satellite-based remote sensing, such as temporal, spatial, and spectral resolutions potentially obstruct the prompt identification of pest damage, particularly during the initial stages of infestation (Luo et al., 2023; Mngadi et al., 2024). In contrast, Unmanned Aerial Vehicles (UAVs) excel at capturing within-stand variations and in making single-tree assessments are possible, which enables the tracking of single trees dynamics at a high spatial resolution monitoring of up to millimetre level resolution (Dainelli et al., 2021). For these reasons, an array of research endeavours has embraced UAV technology to assess diseases and defoliation within forestry environments (Garcia et al., 2023; Lin et al., 2019; Mngadi et al., 2024; Xu et al., 2022). For instance, Liao et al. (2022) used multispectral UAV imagery to discriminate healthy from red leaf dieback *Eucalyptus* trees. The red-edge, near-infrared, and green indices proved most sensitive, and RF and Spectral Angle Mapper were identified as suitable methods (Liao et al., 2022).

Despite these apparent successes and regardless the platform, very few studies have made it to the operational stage. Many studies rely on single datasets for model construction and testing (e.g., Adelabu et al., 2014; dos Santos et al., 2017; Liao et al., 2022; Minařík et al., 2021; Panopoulou et al., 2024). For example, Ma et al. (2023) achieved 94 % accuracy when assessing *Pinus yunnanensis* canopies attacked by *Tomicus* spp. with hyperspectral data, but relied on a single dataset, which was split randomly in training (70 %) and test (30 %) data. The same approach was used by Xu et al. (2022a) when developing a multispectral-based method for monitoring *Pantana phyllostachysae Chao* infestation severity in bamboo forests. Methods relying on a single dataset could be useful when the goal is to upscale a limited number of ground truth measurements to the entire forest stand level. However, in the vast majority of studies, the (often implicit) goal is to develop a stand-alone model, i.e., a model that can be deployed without requiring an extensive training data set. Due to its susceptibility to overfitting, a model developed for a single site may lack the robustness required for broad applicability. Even comparisons between models or feature selection processes should be approached with the highest caution.

The purpose of the current study is to build a robust model that detects *Gonipterus* sp. n. 2 damage levels in *Eucalyptus* stands. Apart from being the first study – to our knowledge – to assess *Gonipterus* sp. n. 2 damage with UAV data, one of the innovative aspects is that data collected at multiple sites and across multiple seasons. The first goal of this study was to examine the effects of damage caused by *Gonipterus* sp. n. 2 on spectral reflectance across several sites. Next, the performance of different machine learning model is compared, including feature selection to check consistency within each model and across models. Models are trained and tested on single sites, on all sites combined and on independent sites not included in the training. Finally, the study evaluates whether a more expensive 10-band sensor is required, or whether a more economical 5-band multispectral sensor can suffice.

2. Methods and materials

2.1. Study area

UAV flight missions were carried out over several *E. dunnii* plantations under the management of SAPPI Forests in Southern Africa (Fig. 1) situated in four areas (Ixopo, Melmoth, Greytown and Piet Retief) in the KwaZulu-Natal and Mpumalanga provinces (Table 1). Based on Koppen-Geiger climate classification topology (Conradie, 2012), Greytown, in the KwaZulu-Natal Midlands, has a monsoon-influenced temperate oceanic climate; Melmoth, in the inland section of Zululand, has a humid subtropical climate; Piet Retief, in the Mpumalanga Province, a warm temperate climate; and Ixopo, in the KwaZulu-Natal Province, a temperate oceanic climate.

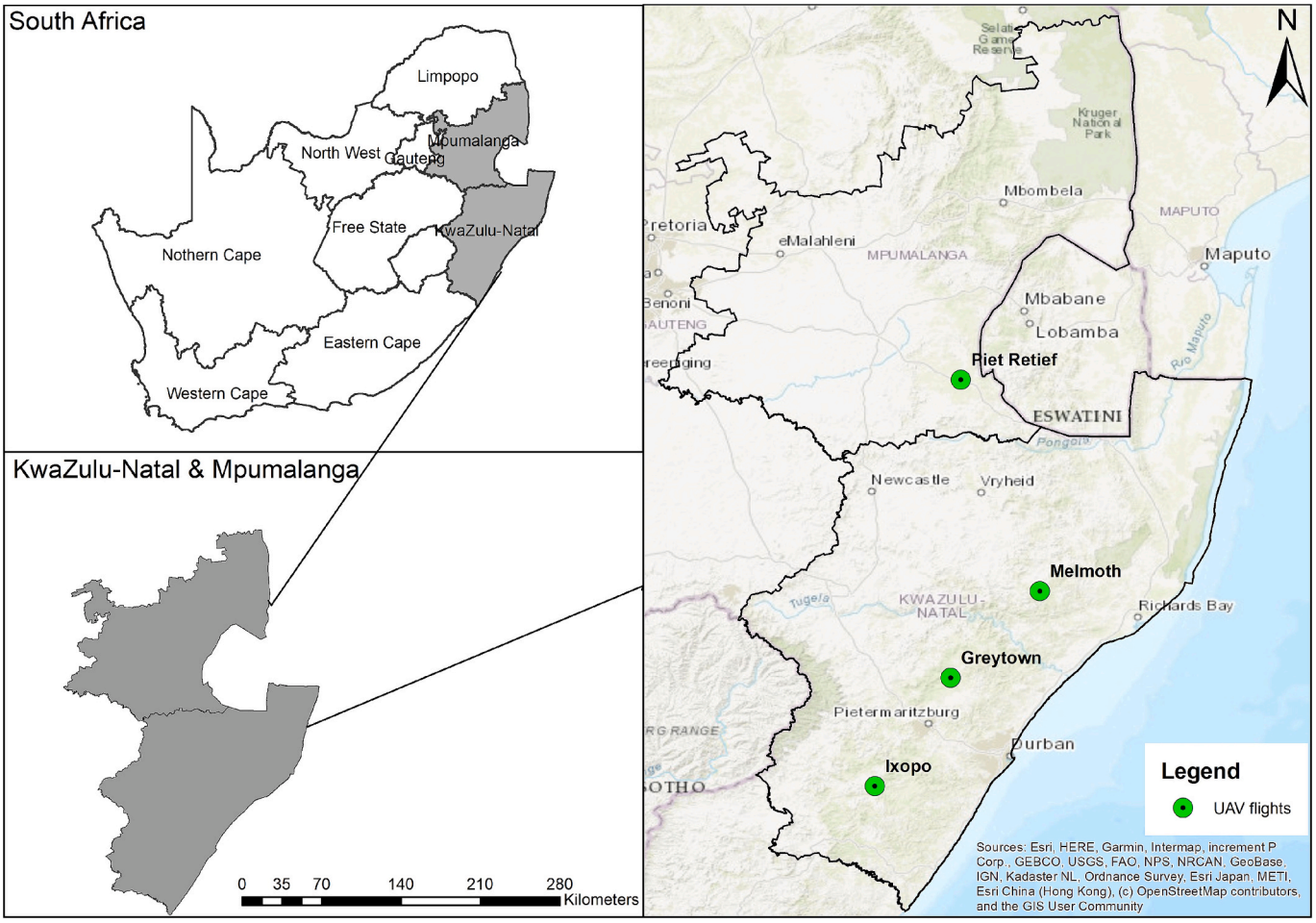


Fig. 1. Location of study areas in KwaZulu-Natal and Mpumalanga, South Africa.

Table 1
General location information of *Eucalyptus dunnii* stands assessed in Mpumalanga and KwaZulu-Natal provinces. All repeated measurements were done over the same area within the stand.

Area	Mean annual rainfall (mm)	Mean annual temperature (°C)	Altitude (m)	Coordinates	No. of UAV flights	Compartment	Planting date
Greytown	817 mm	16.7 °C	1106 m	29.236911°S 30.568441°E	2	Gre	December 2021
Ixopo	742.0 mm	17.5 °C	1210 m	30.093886°S 30.090286°E	3	Ixo	February 2021
Melmoth	817 mm	18.9 °C	1240 m	28.560344°S 31.19085°E	1	Mel	November 2020
Piet Retief	850.0 mm	16.6 °C	1251 m	26.866136°S 30.646256°E	1	PR1	March 2020
Piet Retief	850.0 mm	16.6 °C	1251 m	26.866136°S 30.646256°E	1	PR2	March 2020
Piet Retief	850 mm	16.6 °C	14,765 m	26.910099°S 30.582280°E	1	PR3	March 2020

2.2. Field data collection

In total, 9 flight campaigns were carried out between February 2022 and March 2023 (Table 2). The selection of plantation stands was based on age and species (juvenile *E. dunnii* stands between 1 and 2 years old, Table 1) and guided by reports from foresters concerning high infestation levels of *Gonipterus* sp. n. 2. In Ixopo and Greytown, two experimental trials were utilized, in which half of the plots had been treated with an insecticide to prevent damage by *Gonipterus* sp. n. 2. Each trial site is 3.25 ha and consists of six block plots, three treated and three untreated plots (Fig. A1). Each plot contains 900 trees (30 × 30). At Melmoth and Piet Retief, ad hoc flight campaigns were done in commercial plantations following reports of high infestation levels.

Each UAV flight campaign was accompanied by a field sampling campaign, where a visual assessment was conducted of the damage level of 89–100 trees within the stand (Fig. 3). Trees were selected through stratified random sampling ensuring sufficient trees of different damage levels are present in each stand. Four different damage classes were discerned based on the affected tree canopy (Fig. 2): (a) No, (b) Low (1–25 %), (c) Medium (26–70 %), and (d) High damage (>70 %), based on both foliage and leaf discolouration in the top of the canopy (Fig. 2). The assessment focused on the new flush of the tree canopy. The assessors received training from an entomologist on how to assess the damage level through visual assessment and consensus assessment was

Table 2
UAV flight information.

Area and stand	Time and date of flight	Flight altitude (m)	Ground sampling distance (cm)
Gre_Feb22	01/02/22 11 h30	80	5.5
Gre_May22	04/05/22 12 h30	80	5.5
Ixo_Feb22	09/02/22 11 h00	80	5.5
Ixo_Apr22	14/04/22 11 h00	80	5.5
Ixo_May22	05/05/22 12 h00	80	5.5
Mel_Jan23	31/01/23 12 h20	30	2.0
PR1_Mar23	01/03/23 11:45	30	2.0
PR2_Mar23	02/03/23 12 h15	30	2.0
PR3_Mar23	02/03/23 13 h00	30	2.0

then conducted among the assessors to ensure consistency. Furthermore, reference images of *Gonipterus* sp. n. 2 damage levels of each class were used as a guideline to discern the different classes Fig. 2.

The precise location of each sampled tree was recorded with a precision Global Navigation Satellite Systems (GNSS) device (S10 GNSS Receiver, Stonex, Paderno Dugnano, Milan, Italy). In cases where the Real-time kinematic (RTK) network was inaccessible, we resorted to the nearest trig beacon within a 3-km radius of the flight area, which served as base station.

2.3. UAV flights

The UAV flights were conducted using a DJI M300 UAV (DJI, Shenzhen, Guangdong, China), equipped with a 10-band MicaSense RedEdge Dual MX Camera (AgEagle, Wichita, KS, USA) was available (Table 2). In each stand, four permanent ground control points (GCPs) were placed at the corners of the stand and their location and altitude were measured with the RTK-GNSS. The DJI Pilot 2 app was used for waypoint flight planning and execution. Flights were conducted at altitudes of 30–80 m (Table 2), with those before May 2022 flown at 80 m and later flights adjusted to 30 m. A horizontal and vertical overlap was set at 80 % Maes (2025), with the flight speed set to 5 m/s was maintained. Before each flight, a single grey calibrated reflectance panel was captured by holding the drone about 1 m above the panel, avoiding

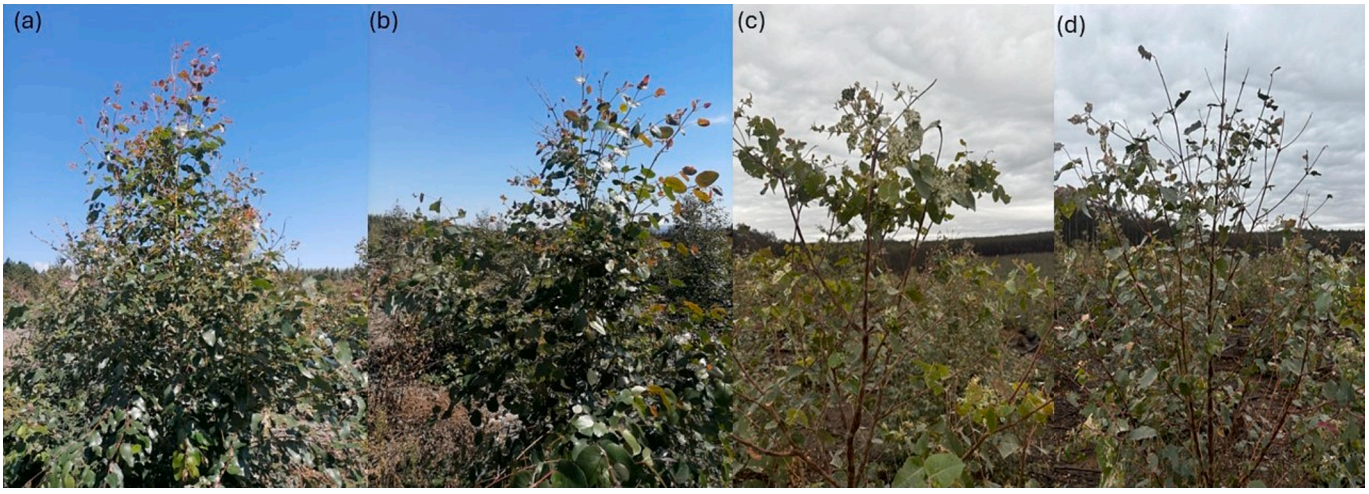


Fig. 2. *Gonipterus* sp. n. 2 defoliation severity classes. (a) No damage, (b) Low 1–25 % damage, (c) Medium 26–70 %, (d) High >70 %.

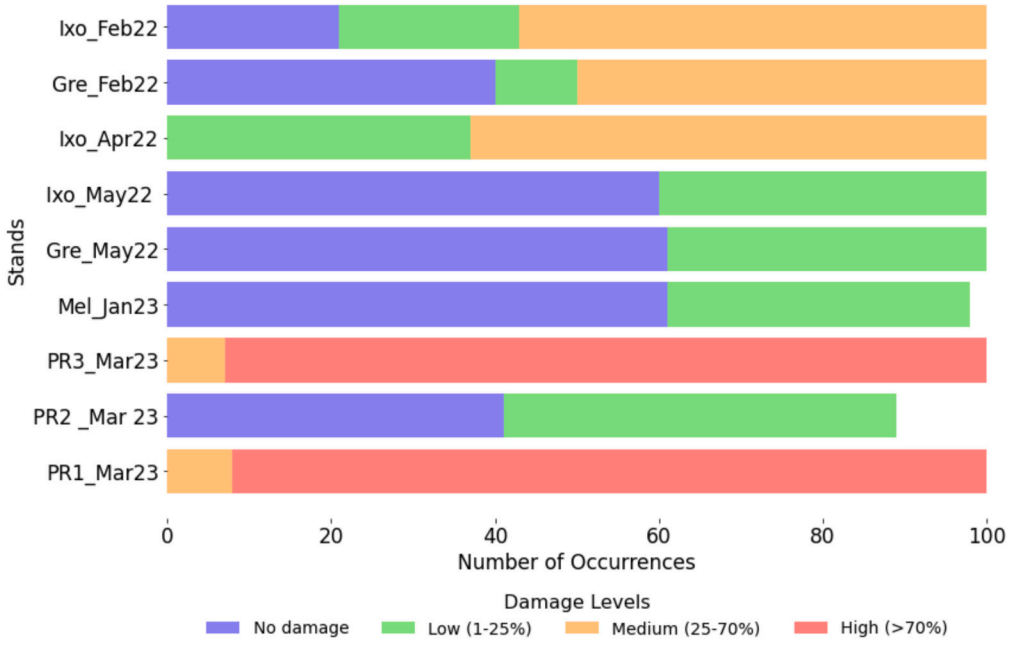


Fig. 3. Distribution of sampled damage severity classes across compartments.

shadows on the reference panel or on the downwelling light sensor (Daniels et al., 2023). The flight operations were executed in sunny conditions (Table 2).

2.4. Image processing

Orthomosaic creation was carried out using Agisoft Metashape Professional (version 1.4.3, 64-bit, Agisoft LLC, St. Petersburg, Russia) without prior pre-processing of the imagery. Agisoft Metashape was also used to convert digital numbers (DN) to reflectance, based on the pre-flight grey panel calibration. The downwelling light sensor data was not used given that flights took place in clear sky conditions, where its inclusion in data processing can deteriorate reflectance estimates (Daniels et al., 2023). Ground control points (GCPs) were used for georeferencing, and a digital elevation model (DEM) and digital terrain model (DTM) were generated. The DTM was created by using the *classify ground points* function in Metashape, with default parameter settings, to identify the ground points based on height. The orthomosaic, DEM and DTM were generated and exported from Agisoft Metashape as TIF files for further processing. The exported orthomosaic, which represents radiometrically calibrated data, was normalized to a 0–1 range in QGIS 3.30 using the raster calculator to ensure consistency in pixel value across datasets. To harmonize spatial resolution for cross-dataset comparison Borra-Serrano et al. (2015), the 2.0 cm data from the 30 m flights were down sampled to 5.5 cm using the nearest neighbour method in QGIS 3.30.

Histogram matching was conducted for all UAV flights to ensure consistency across different datasets. Histogram matching uses the source and the reference image's cumulative distribution function (CDF) and can be used to normalize data collected at different locations and measurement conditions (Lauwers et al., 2024). The Gre_Feb22 dataset used as reference. For each pixel reflectance value in the source image, the algorithm finds the corresponding pixel reflectance value in the reference image with a similar CDF (Helmer and Ruefenacht, 2005). This mapping of pixel intensities transforms the source image histogram to match the reference image histogram.

2.5. Tree crown classification

The vegetation height was calculated as the difference between DEM and DTM and used to delineate the tree canopy by selecting all pixels with a vegetation height > 0.5 m. The first delineation did not achieve the desired level of accuracy to isolate pure crown pixels. In a second step, we therefore performed a further refinement to remove soil or mixed pixels from the initially selected pixels by employing a supervised RF segmentation algorithm (python scikit-learn package using libraries scikit-image, RF classifier, numpy and matplotlib) (Mahapatra, 2014). This was performed for each dataset (orthomosaic) separately. The RF model was trained on two classes (tree canopy and background). A visual inspection was conducted for each stand to ensure satisfactory results, confirming that no grass or soil canopy was present in the segmentation. No accuracy assessment was performed.

2.5.1. Vegetation indices retrieval

Twenty-nine vegetation indices were computed in Python 3.7 to be included in the model (Table A1) (Python Software Foundation, 1991). These indices have proven to be effective in characterizing the extent of vegetation defoliation (leaf biomass) and chlorophyll content (Estrada et al., 2023).

2.5.2. Texture feature extraction

We extracted canopy structural properties with the Grey Level Co-occurrence Matrix (GLCM) calculated in Python using the greycomatrix and greycoprops methods in scikit-image. We employed second-order texture measures to detect finer-scale variability while considering spatial relationships among neighbouring pixels. Previous studies have demonstrated that a 3×3 sliding window is sensitive to fine scale variation in pixel brightness compared to larger window sizes (Kelsey and Neff, 2014), so we used this window size. The mean texture values were extracted using the zonal statistics tool in python which coincided with the field samples. The GLCM incorporates a range of filters, including contrast, variance, dissimilarity, mean, homogeneity, correlation, entropy, and second moment, which were computed for each spectral band (Table A2). This resulted in a total of 80 GLCM output variables, see Table A2 for an overview.

2.5.3. Feature retrieval

The sampled trees were identified using the precision Global Navigation Satellite System (GNSS) information. A circular polygon with a 0.5 m diameter from the central stem position was measured using GNSS and generated for further analysis. Zonal statistics extracted the mean spectral reflectance value of all vegetation indices and texture features for each tree.

2.5.4. Statistical analysis

A Type III Analysis of Variance (ANOVA) with damage class ($n = 4$ levels) as the main factor, and dataset as nested variable was used to evaluate the performance of the vegetation indices. A Tukey honest significant difference was used as post-hoc test. Prior to the ANOVA, Levene's test was used to assess homoscedasticity. When violations were detected, various transformations were tested (log, square root, reciprocal, box cox and power transformation), and the best performing applied. In case this was not successful, Kruskal Wallis KW (with damage class as single factor) was used instead, followed by Dunn's post-hoc test.

2.5.5. Machine learning models and recursive feature elimination

A machine learning model was developed for assessing the *Gonipterus* sp. n. 2 damage levels. We included three common machine learning models, namely Support Vector Machine (SVM), Random Forest (RF) and Extreme Gradient Boosting (XGBoost). Several input models were tested: 1) using spectral reflectance bands; 2) using only vegetation indices ($n = 25$ and $n = 29$) for the five and 10 band version; 3) using texture data; 4) merging spectral indices and bands; 5) using a combination of spectral bands, vegetation indices and texture data. To select the best model, all the datasets from all fields were merged.

The machine learning models were implemented in Python using Scikit-learn with hyperparameter tuning conducted for all three machine learning models using Bayesian Optimization. For XGBoost, the search space included learning rate (0.01–0.3, log-uniform), maximum tree depth (3−10), number of estimators (100−1000), subsample ratio (0.5–1.0), column sampling by tree (0.5–1.0), and gamma (0–5). The RF parameters included the number of estimators (100–1000), max_depth (3−30), max_features (0.1–1.0), min_samples_split (2−20), min_samples_leaf (1−20), and bootstrap (True/False). For SVM, the grid explored regularization parameter C (log-uniform: 1e-4 to 1e4), tolerance tol (log-uniform: 1e-4 to 1e-2), and max_iter (100,000 or 500,000), with class weight fixed to 'balanced' to handle class imbalance.

To ensure robust testing and evaluation, the model dataset was divided into a training set (70 %) and a test set (30 %) in the balanced dataset, where each stand was similarly represented with five model iterations. In the field, reference data collected on the damage levels were not balanced per stand, since not every damage class level was present in every stand at the time of monitoring. This inevitably resulted in an imbalanced dataset (Table 4), which could potentially lead to inherent biases in model performance (Chabalala et al., 2023; Devi et al., 2019; Quan et al., 2021). To overcome this, random under sampling was performed using the Imbalanced-Learn Python package, thus reducing the observations in the oversampled classes to match the samples of the class with the fewest samples, resulting in a balanced dataset. The synthetic minority oversampling technique (SMOTE) was tested and showed similar results as random under-sampling method.

The importance of multicollinearity and overfitting was addressed by implementing recursive feature elimination, using a 5-fold stratified cross-validation to select the most important features. Gini importance was calculated to evaluate feature importance in RF and XGBoost models. For the SVM model, a variance threshold was applied to remove low-variance features before training, and feature selection was performed using LinearSVC coefficients. An overview of the amount of training and test data for each model is given in (Table 4).

In addition, the performance of the different model scenarios was also compared for the same input models, but then using only the spectra as would be available for the standard 5-band camera (see Table 3); here, 25 instead of 29 vegetation indices were included for the vegetation index-only and combined models.

Once the overall best performing method was established, the following models were additionally constructed using this method:

- 'Upscaling' scenario: Model trained and tested per site, using a random 70/30 split
- Leave-one-stand-out: Model was trained on all but one stand and tested on the remaining stand to determine its robustness on unseen data.

3. Results

3.1. Distribution of damage classes across compartments and average spectral reflectance

Increasing levels of *Gonipterus* sp. n. 2 damage resulted in lower levels of reflectance in both the visual and the NIR spectrum Fig. 4. This pattern was observed consistently across all stands (with the exceptions of PR1_Mar23 & PRI_Mar23, Fig. 4g & i), although the absolute values of the reflectance differed. The spectral reflectance pattern had minor changes after histogram matching but had a similar trend to the patterns obtained without matching (Fig. A2).

3.2. Spectral vegetation indices across sites

Of the 29 vegetation indices, 21 were analysed with Kruskal-Wallis, as they did not meet the Levene hypothesis of equal variances after the data transformation (Table A3). The overall results showed significant effects for damage, site and their interaction (Damage x Site) especially for NDRE740, NDVI, NGRE and SIPI (Table A3). Overall, a similar trend in estimated marginal means emerged, but the absolute values of vegetation indices varied across stands, even after histogram matching (Fig. 5).

3.3. Modelling Gonipterus sp. n. 2 damage levels

3.3.1. Model comparison and overall performance

We used the merged dataset of all sites, with a random split, to evaluate the performance of the different models. Overall, XGBoost performed best, slightly better than RF, and significantly better than SVM. We therefore focus the data in this section on XGBoost output; the results for RF and SVM were very similar and are given in (Table A6,8). Models using histogram matching performed clearly better than uncorrected datasets (Table A6 and Table A8); therefore, we also only include data from histogram matched data. Overall, the model using VIs, spectra and texture data as input clearly performed best, both when 10-band data were used as input (Table 5). The added value of including texture data is limited, but not negligible. The confusion matrices of the combined input for both the 10-and reduced 5-band datasets showed a slight difference (Fig. 6). With the 10-band sensor, "Low" damage levels were not always distinguished correctly. This trend persisted in the reduced 5-band sensor, but misclassifications was more common.

Table 3
MicaSense ten band multispectral band properties used in this study. Bold items represent classical five band multispectral camera.

Band	Bandwidth
Coastal blue 444 nm	28 nm
Blue 475 nm	**32 nm**
Green 531 nm	14 nm
Green 560 nm	**27 nm**
Red 650 nm	16 nm
Red 668 nm	**14 nm**
Red-edge 705 nm	10 nm
Red-edge 717 nm	**12 nm**
Red-edge 740 nm	18 nm
Near infra-red 842 nm	**57 nm**

3.3.2. Feature importance

Fig. 7 shows the feature weights for 5 runs of the XGBoost, RF and SVM models (all using Histogram matching) for the model having VIs+Spectra as input - the VIs+Spectra+Texture model has 119 variables in the 10-band version, which makes it more difficult to keep the overview. XGBoost and RF show good consistency both between different model runs, as between both models, particularly in the 10-band model feature selection. In the 10-band model, wavelengths at ρ_{444}, ρ_{560} and ρ_{650} had the highest weight, together with PRI, SIPI and, for XGBoost only, CI. In the five-band data, reflectance in the red spectrum (ρ_{668}) is important, next to ARI, CLGREEN and, for XGBoost only, CI. The feature weights for SVM vary significantly across the five iterations in the 10-band model but are more stable in the reduced 5-band model. This pattern differs from the trends observed in XGBoost and RF. SVM consistently prioritizes vegetation indices in its feature selection.

3.3.3. Performance on individual sites

This overall model scenario could generalise well across stands, with overall model accuracies ranging from 76 % to 99 % between stands (Table 6). Models trained and tested on single stands (XGBoost, VIs+-Spectra+Texture as input, random 70/30 train/test split) had still slightly higher accuracies, ranging between 82 % to 100 % (Table 6). Fig. 8 shows the predicted damage levels at stand level for the Ixo_Feb22 stand for all three modelling scenarios and its spatial distribution.

The leave-one-stand-out validation approach performed poorly across all stands, with overall accuracies ranging from 0 % to 68 % (Table 6).

4. Discussion

4.1. Effect of Gonipterus sp. n. 2 damage on leaf spectral reflectance and vegetation indices

This study aimed to assess and model *Eucalyptus dunnii* canopy damage caused by *Gonipterus* sp. n. 2 across multiple stands in South Africa through UAV-based multispectral imaging. Where previous studies collected data at plot level and worked with single-date satellite data, we assessed the ground truth at individual tree level and collected data from different sites and at different stages in the growing season. Despite this diversity in sampling, the study showed consistent patterns of spectral reflectance affected by damage, with decreases in reflectance across the entire spectrum with increasing levels of damage.

The decrease in NIR was largely expected and aligns with a decrease in leaf area index and chlorophyll content (Daughtry et al., 2000; Haboudane et al., 2002; Kimura et al., 2004). However, the reduction in the visible bands was unexpected, as we anticipated a decrease in chlorophyll content for damaged trees, which would normally result in an increase in (particularly) the red and green spectrum. In our opinion, the decrease in the visual spectrum is due to the fact that *Gonipterus* sp. n. 2 larvae prefer to feed on new leaf shoots, which tend to be brighter,

Table 4
Model training and test samples used to develop the machine learning models based on a balanced dataset merging all variables.

Models developed	Training	Test	No of variables	Original unbalanced damage levels			
				No damage	Low	Medium	High
10-band model	518	222	119	284	233	185	185
Reduced 5-band model	518	222	70	284	233	185	185

Table 5
Overall accuracies for each model for the 5- and 10-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) after histogram matching using XGBoost classifier.

Models	10-band	5-band
Spectra	84 %	80 %
VIs	82 %	81 %
Texture	76 %	76 %
VIs+Spectra	83 %	80 %
VIs+Spectra+Texture	92 %	88 %

Footnote: The Reduced 5-band used the 5 spectral bands from the Red edge MX 10.

Table 6
Overall site performance of XGboost classifier on trained model per site, overall model and leave e site out model accuracies.

Site	Trained per site	Overall model	Leave one site out
Gre_Feb22	92 %	95 %	40 %
Gre_May22	99 %	76 %	39 %
Ixo_Feb22	82 %	85 %	35 %
Ixo_Apr22	99 %	93 %	48 %
Ixo_May22	98 %	97 %	68 %
Mel_Jan23	98 %	86 %	43 %
PR1_Mar23	99 %	97 %	7 %
PR2_Mar23	100 %	99 %	9 %
PR3_Mar23	94 %	94 %	0%

thereby exposing branches of older leaves. These older leaves typically have higher pigment concentration (Kume, 2017) and reflect less light in the visual spectrum. However, this hypothesis requires confirmation through future research, in field or greenhouse trials.

In the ANOVA, consistent differences were found in NDRE740, NGRE, and SIPI, showing significant effects for Damage, site and their interaction. The post hoc test confirmed clear separation among damage levels, while CI, ARI, OSAVI and MTCI showed no significant differences. Month was not included as a covariate in the model, as data collection was conducted on an ad hoc basis following reported damage events, and the limited number of temporal replicates precluded a reliable assessment of time-related effects. Vegetation indices such as SIPI, ARI, PRI and CI were selected as important features in the model development for either five- or 10-band multispectral cameras in the XGboost classifier. CI is a vital vegetation index in determining chlorophyll content, while ARI is closely related to leaf anthocyanin, known to accumulate in response to stress for photoprotection (Mielke et al., 2012; Stone et al., 2001). This aligns with Lottering et al. (2019) who detected the effects of *Gonipterus* sp. n. 2 on chlorophyll content and selected ARI as important feature using high resolution satellite data. PRI was also singled out as an important variable, and can here be linked to the ratio of carotenes over chlorophyll ratio (Lin et al., 2023; Liu et al., 2020; Moley et al., 2022). Nevertheless, despite consistent trends of damage levels on VIs, absolute VI levels were different across different status, with far-reaching consequences on the Machine Learning models (See further, S 4.2).

4.2. Machine learning model performance

We tested different ways for data pre-processing and different inputs for the models. Histogram matching, which was used to correct for seasonality and site-specificity in other studies (Lauwers et al., 2024), was partially successful in normalising the reflectance values across multiple flight campaigns, with a clear improvement of the models that used histogram matching (Tables 5, A6). However, histogram matching was clearly not able to correct for all site differences (Fig. 4) and was not sufficient for creating a robust model performing well on untrained data (Fig. 6). The Pseudo-invariant feature radiometric correction method is a potential alternative to histogram matching and could produce better results than histogram matching (Xu et al., 2021).

While models using only GLCM texture data consistently yielded the lowest accuracies, adding texture to models with spectra and vegetation indices improved the prediction of *Gonipterus* sp. n. 2 damage. This complementary role aligns with Lottering et al. (2019), who demonstrated added value using WorldView-2 imagery, though the significant GSD difference (1.8 vs 0.02 m) underscores the critical impact of resolution on texture variables. Still, if only texture was used, the model performance was relatively weak, and the contribution of texture features to the overall model relatively limited. We used a 3 × 3 GLCM window and had a GSD (spatial resolution) of 5.5 cm raising questions about the role of these specific parameters. Notably, while reducing the ground sampling distance is often assumed to enhance texture-based classification, our results showed that individual site models using higher-resolution data did not perform better than those based on 5.5 cm resolution (not shown). Hence, increasing spatial resolution alone did not improve texture-based detection. We initially also tested multiple window options (3 × 3, 5 × 5, and 7 × 7), and the 3 × 3 window yielded the best performance, likely reflecting the fine-scale, localized nature of early damage. We therefore focused on the 3 × 3 window in this work. Future work could explore other texture descriptors or multi-scale approaches to better capture texture metrics' sensitivity to subtle canopy changes.

Our results highlight XGBoost as the top-performing model in terms of both classification accuracy and feature selection. This finding is consistent with previous research (Huang et al., 2022; Kasinathan and Uyyala, 2021; Xu et al., 2021). Hennessy et al. (2020) cautioned that feature selection can be inconsistent both within several repetitions of the same method and across methods, and the selected features should therefore be interpreted with care. However, in our study, XGBoost showed good consistency across the different runs. Moreover, RF also showed similar consistency, with a very high overlap in features between XGBoost and RF, suggesting the stability and robustness of the selected features. SVM, on the other hand, performed clearly worse, and was less consistent in the selected features across different runs.

Trees with medium and high damage levels were reliably distinguished from each other and from the other damage classes, but the 10-band model struggled a bit to classify the Low damage correctly. However, it must be noted that even in the fieldwork for the damage class acquisition, there was least consistency across the different assessors for the Low damage class, so some error in the ground truth data might be present. From an application perspective, getting accurate classification of the Low damage class is probably also least relevant, since low damage levels generally do not lead to measurable growth impacts.

The models based on the 10-band multispectral camera performed

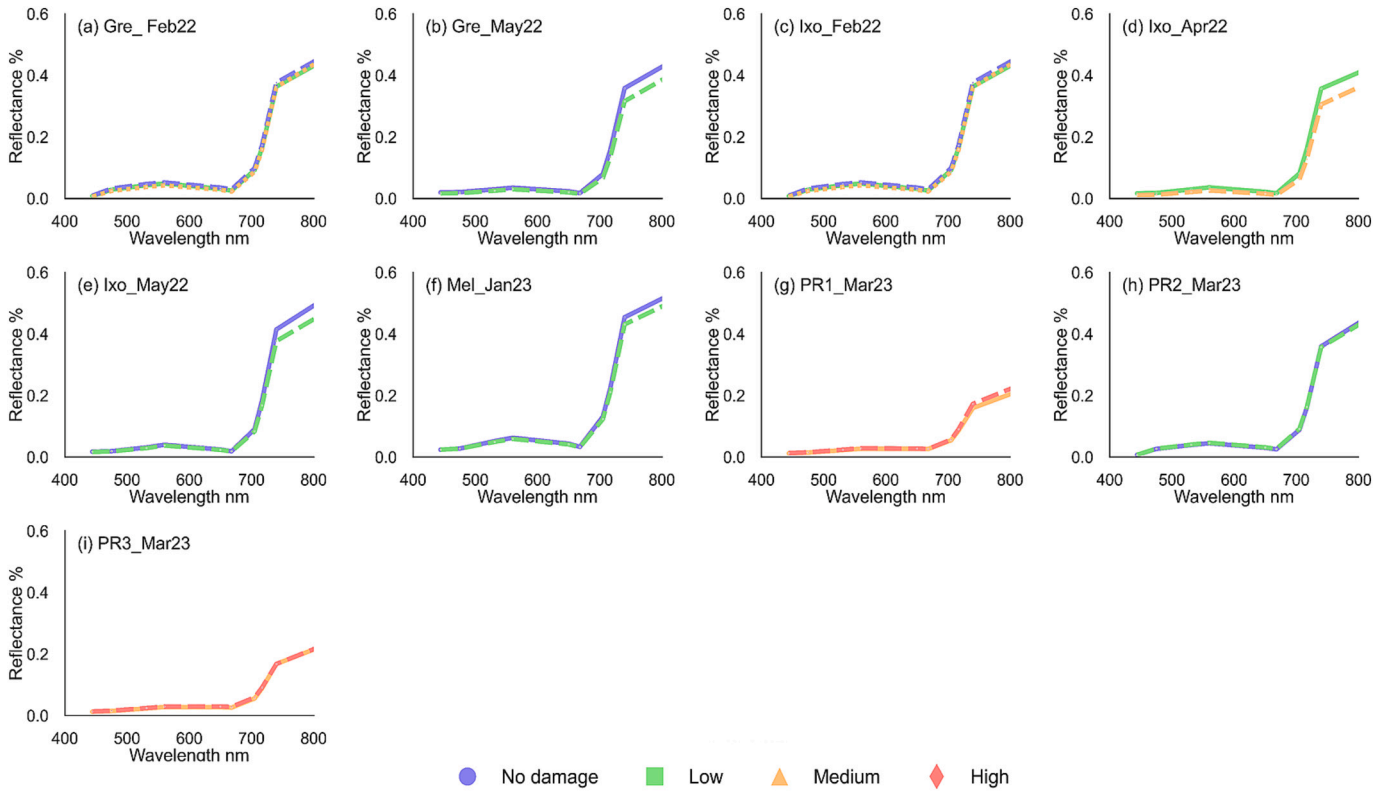


Fig. 4. Average spectral reflectance across multiple sites based on five and ten band cameras and the level of *Gonipterus* sp. n. 2 damage after histogram matching.

better than those with the five-band spectra, although the difference was relatively small (92 vs 88 %). Note that the 10-band model manages to distinguish the Medium damage better than the 5-band model, which can be quite relevant for the applications. Overall, the 10-band multispectral camera is clearly to be preferred, but in case financial restrictions impede its use, measurements with the 5-band camera can still give reliable estimates.

4.3. Performance at stand level

We found excellent performance for models trained on individual sites, with accuracies between 82 and 100 % (Table 6), illustrating that the method is highly suited for upscaling studies, as illustrated in Fig. 8. Numerous other UAV studies focusing on pest and disease detection have used single date imagery or time series of one single location or very few sites (Lottering et al., 2019; Nguyen et al., 2023; Pan et al., 2023; Wu et al., 2023). In fact, the work done on *Gonipterus* sp. n. 2 detection (Lottering et al., 2020, 2019; Lottering and Mutanga, 2016) all used single site datasets, with models trained and tested on one single satellite image.

The overall model approach tested on each site also performs well in the individual stands (Table 6), demonstrating strong generalization and effectiveness in consistently detecting *Gonipterus* sp. n. 2 damage at each stand. However, the leave-one stand-out validation method did not give satisfying results, even after data normalization through histogram matching. The poor performance can have several causes. The imbalance in the dataset probably played a role. Finding sites with *Gonipterus* sp. n. 2 damage was challenging due to the sporadic nature of outbreaks and the rapid recovery of *Eucalyptus* trees post-rainfall (Chambi-Legoas et al., 2022). Encountering all four classes of *Gonipterus* sp. n. 2 damage levels within a single compartment proved particularly challenging, leading to unbalanced datasets per site for model development. In conclusion, a larger and more balanced dataset and additional data collections are needed to improve generalization and support transfer learning.

The Leave One Site Out model clearly did not perform well. Several improvements were tested, but without much success. Most importantly, seasonality and phenology were not incorporated in the model, but obviously influence plant properties and canopy reflectance significantly. To address this, we attempted to normalize UAV-derived vegetation indices using Sentinel-2 data, adjusting for regional and temporal patterns observed in satellite-based indices. However, this correction did not yield consistent performance improvements. Consequently, the normalization approach was excluded from our final models. Another option would be to incorporate phenological stage data observed at the site in the model (such as detailed growth stage, flowering, or leaf area index measurements). Collecting such information concurrently with UAV flights proved infeasible due to resource and logistical constraints. Satellite-derived LAI offered a potential proxy, but its temporal coverage and resolution limitations, as highlighted above, rendered it impractical for our specific study design. Future work should prioritize multi-seasonal datasets or develop robust phenological correction strategies, particularly where natural phenology might obscure defoliation signals.

In addition to seasonality, further improvements could include more strict data capture and improved processing protocols. Possible improvements include the implementation of the empirical line method based on multiple grey reference panels Daniels et al. (2023) manual setting of the multispectral camera Bagnall et al., 2023; Swaminathan et al. (2024), setting the flight direction relative to the solar azimuth and adjusting the time of flight to have a constant solar zenith angle across datasets Maes (2025) and/or corrections for sun-sensor geometry (Bidirectional Reflectance Distribution Function, BRDF) (Heim et al., 2024; Jafarbiglu and Pourreza, 2023). Furthermore, additional data collection is required to develop transfer learning algorithms to enable multi-site model generalization.

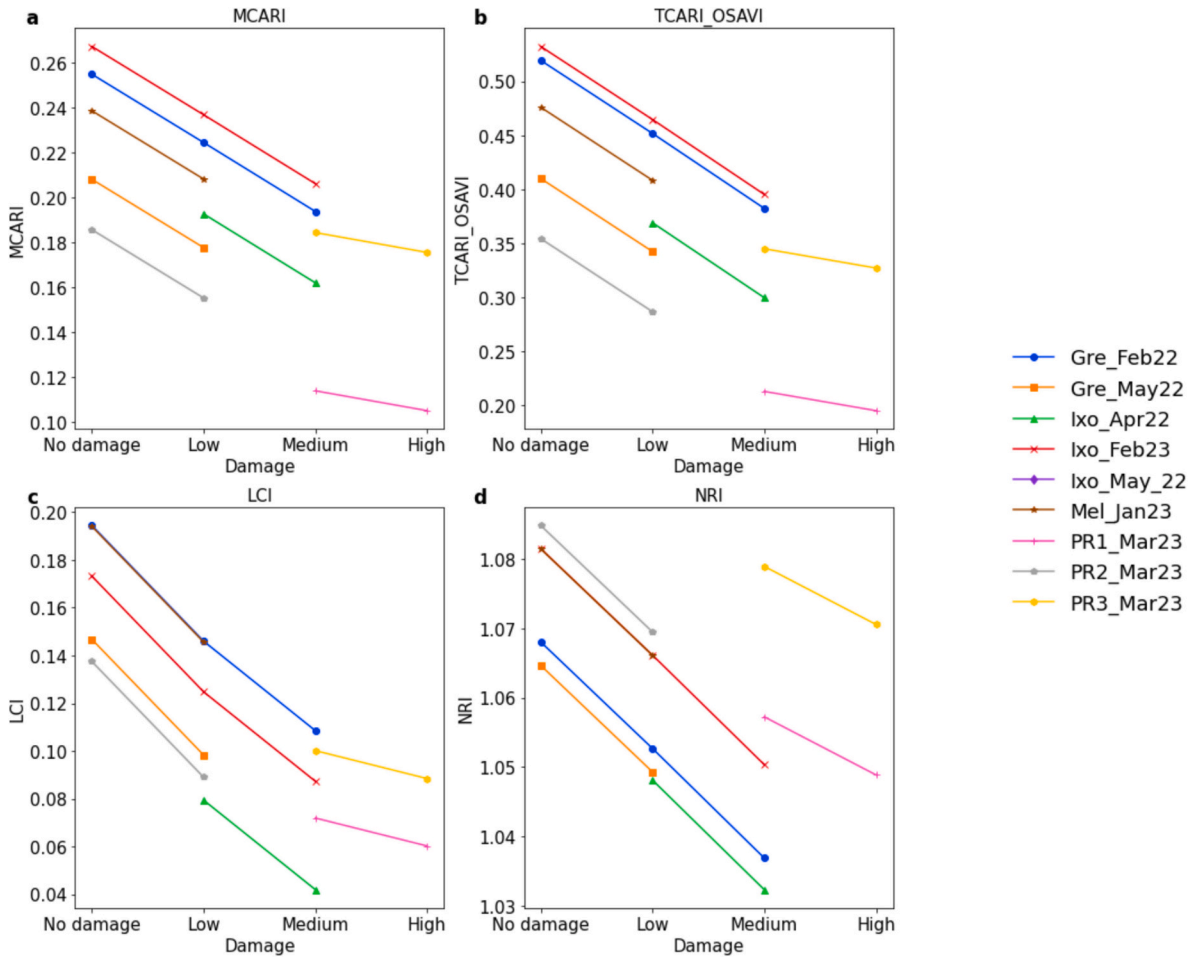


Fig. 5. Estimated marginal means of selected vegetation indices between site and damage across all sites after histogram matching (a) Modified Chlorophyll Absorption Ratio Index (MCARI), (b) Transformed Chlorophyll absorption and Soil Adjusted Vegetation Index (TCARI-OSAVI), (c) Leaf Chlorophyll Absorption Index (LCI), (d) Normalized Redness Intensity Index (NRI).

4.4. *Limitations of the current study*

We cannot guarantee that our model is specific to *Gonipterus* sp. n. 2 damage. We indeed did not have other UAV datasets of different pests and diseases at our disposal and could not test whether the model would wrongfully detect damage for those situations. No other disturbances were present or reported during our field campaign. However, this model is still relevant, given *Gonipterus* sp. n. 2 is one of the most

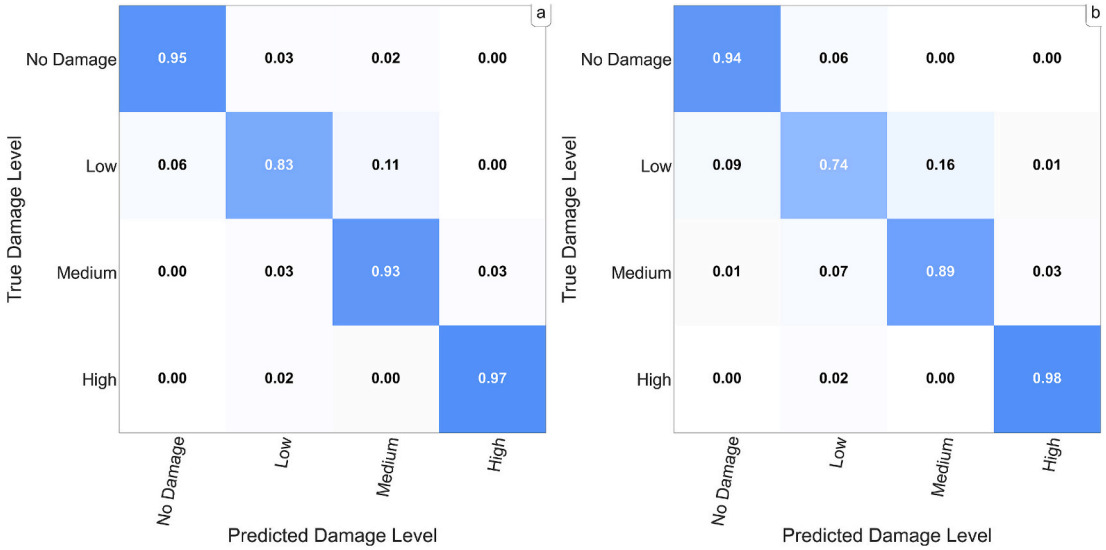


Fig. 6. Confusion matrices of the full datasets, using VIs + Spectra+ Texture as input, for (a) the 10-band data, (b) the 5-band data.

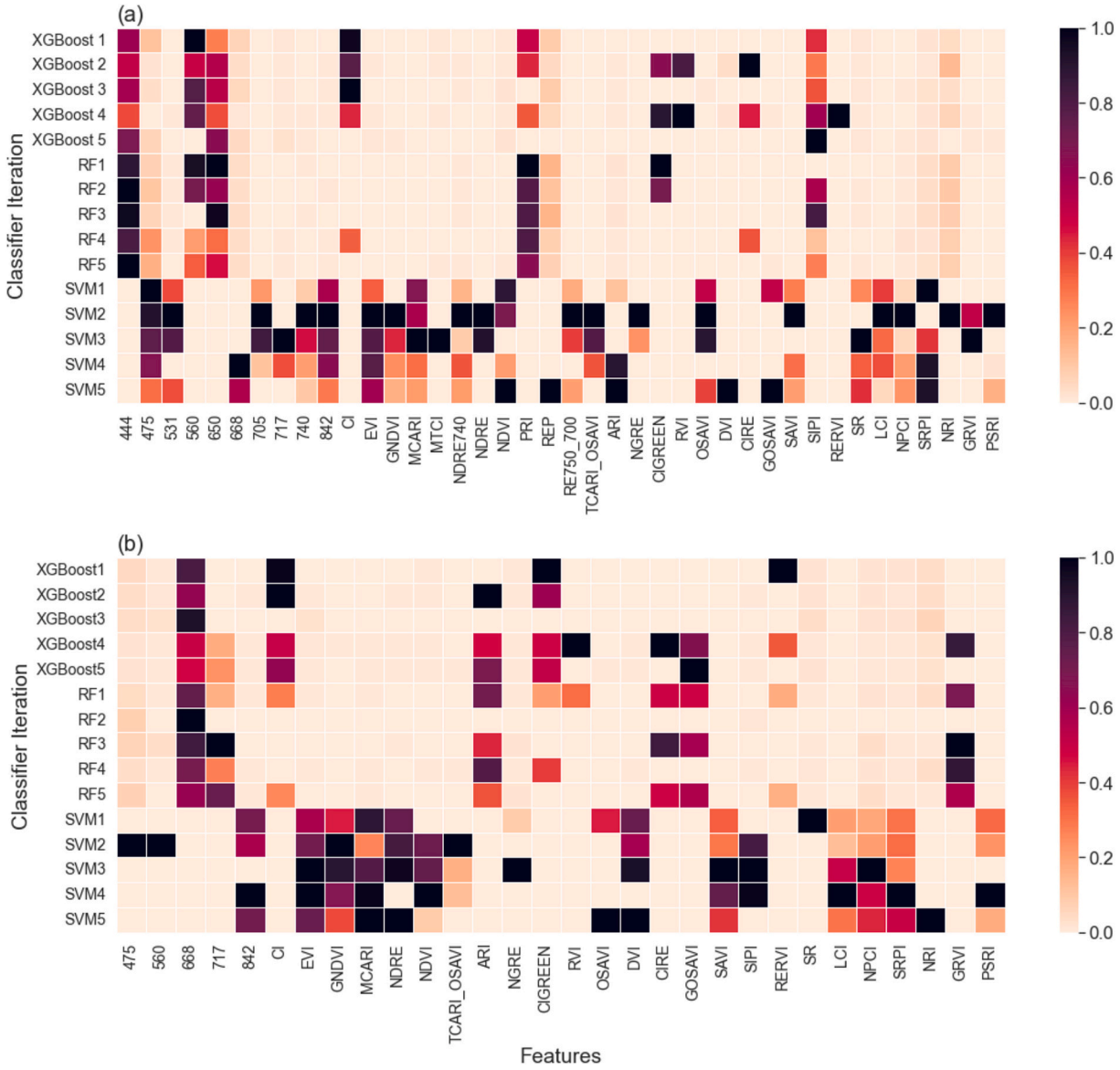


Fig. 7. Selected feature importance, (a) VIs + spectral 10-band data (Micasense-MX Red Edge Dual) camera, (b) VIs + spectral reduced 5 band camera after histogram matching based on five iterations of machine learning classifiers RF, XGBoost and SVM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

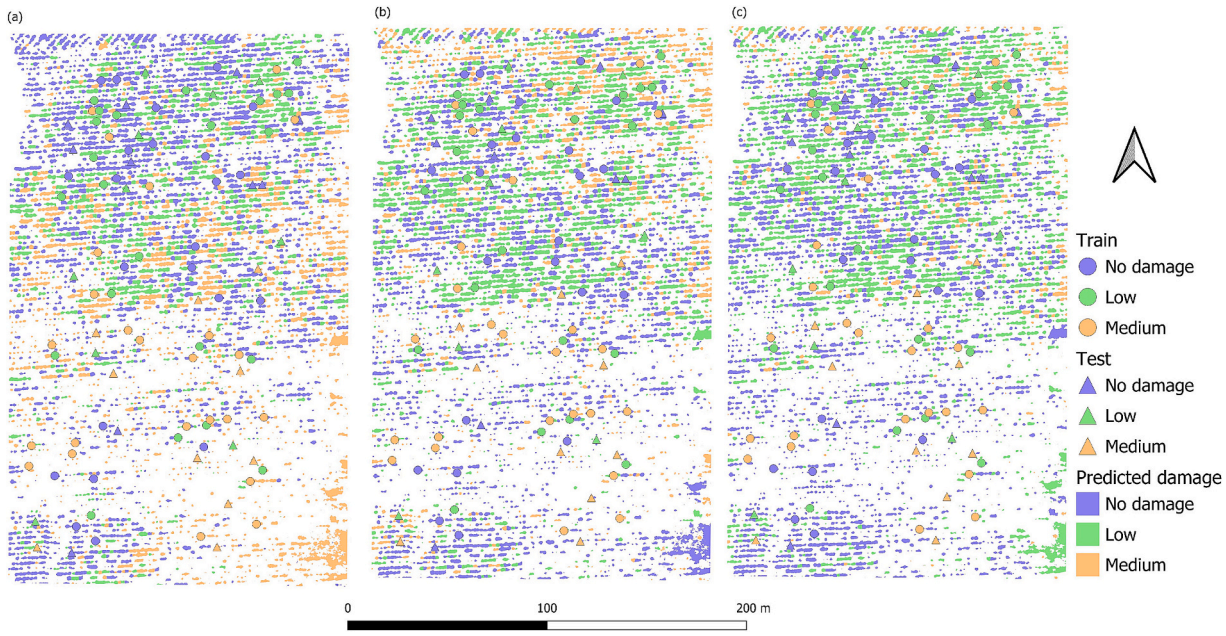


Fig. 8. Spatial distribution of predicted *Gonipterus* sp. n. 2 damage in Ixo_Feb22 based on (a) model trained per site, (b) Overall model, (c) Leave one site out.

important pests, and given the expertise of local foresters. Flights can be targeted when there is local knowledge that *Gonipterus* sp. n. 2 is present to some extent, where the method informs about the severity and extent of the infestation at stand level. Further, as the feeding patterns of other important pest insects of *Eucalyptus* in South Africa, such as *Leptocybe invasa* (Hymenoptera: Eulophidae) or sap-sucking insects Wingfield et al. (2008) are different from those of *Gonipterus* sp. n. 2, we also expect distinct patterns in spectral reflectance. Nevertheless, we highly recommend obtaining remote sensing data as well as ground truth information in a standardized way of different diseases and pests in *Eucalyptus* forests, and creating publicly available libraries of these datasets, in order to create and evaluate more specific models.

Finally, we only included young *Eucalyptus* stands in our datasets, in contrast with other *Gonipterus* sp. n. 2 studies focusing on adult stands (Lottering et al., 2019; Lottering and Mutanga, 2016; Xulu et al., 2024). Assessing *Gonipterus* sp. n. 2 damage in young stands is very relevant, because of the damage and dieback of younger trees Health (PLH), E.P. on P, et al., 2018. We did not sample in adult stands, because the visual assessment of damage levels was not reliable or feasible in adult stands since trees are too tall. Although we expect similar patterns due to *Gonipterus* sp. n. 2 damage in adult stands, further research would be required to test our model on adult stands.

5. Conclusions

This study investigated the development of a remote sensing method to estimate damage levels due to *Gonipterus* sp. n. 2 using UAV-based multispectral data. The results have shown that XGBoost can classify *Gonipterus* sp. n. 2 damage levels across multiple stands and regions, and that models that combine reflectance data, vegetation indices and grey level co-occurrence matrix perform best. The combined scenario of the 10-band model had an overall accuracy of 92 %, the 5-band model performed slightly worse with 88 % accuracy after histogram matching. The outcomes of this study can assist foresters in identifying plantations with *Gonipterus* sp. n. 2 damage using UAVs or mapping the spatial variation of the *Gonipterus* sp. n. 2 damage within the stands.

Declaration of generative AI and AI assisted technologies in the writing process

During the preparation of this work, the authors utilized free version of Grammarly to edit the language and enhance readability. Following the use of these tools, the authors thoroughly reviewed and edited the content as needed, and take full responsibility for the final content of the publication.

CRediT authorship contribution statement

Phumlani Nzuza: Writing – review & editing, Writing – original draft, Visualization, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Michelle L. Schröder:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Rene J. Heim:** Writing – review & editing, Methodology. **Louis Daniels:** Writing – review & editing, Formal analysis. **Bernard Slippers:** Writing – review & editing, Methodology, Funding acquisition. **Brett P. Hurley:** Writing – review & editing, Methodology. **Ilaria Germishuizen:** Writing – review & editing, Methodology. **Benice Sivparsad:** Writing – review & editing, Methodology. **Jolanda Roux:** Writing – review & editing, Methodology. **Wouter H. Maes:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

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

Acknowledgements

PN receives funding from Ghent University Special Research Fund (BOF) and from the University of Pretoria, Tree Protection Cooperative Programme (TPCP). We thank SAPPI for their collaboration, access to their plantations to conduct UAV flights and identifying sites with infestations of *Gonipterus* sp. n. 2. We thank Josias Letaoana (FABI), Xolani Colvelle (ICFR), Enos Ngubo (ICFR) and Letho Nzimande (ICFR) for assisting in the assessment and training. We thank Dr. Adedayo Adeleke

from the University of Pretoria, Department of Geography, Geo-informatics and Meteorology for providing training on using the precision GNSS.

Appendix A

Table A1
Overview of the computed spectral vegetation indices.

Vegetation indices	Formula	Band used	Reference
Normalized Difference vegetation index (NDVI)	$\mathrm{NIR} = \frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+\mathrm{R}}$	842 nm, 668 nm	Rouse et al. (1973)
Anthocyanin reflectance index (ARI)	$ARI = \left(\frac{1}{\mathrm{G}}\right) - \left(\frac{1}{\mathrm{RE}}\right)$	560 nm, 717 nm	(Gitelson et al., 2009)
Leaf chlorophyll index (LCI)	$\mathrm{LCI} = \frac{\mathrm{NIR}-\mathrm{RE}}{\mathrm{NIR}+\mathrm{R}}$	842 nm, 668 nm, 717 nm	(Datt, 1999)
Enhanced Vegetation Index (EVI)	$\mathrm{EVI} = 2.5\mathrm{X}\frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+6\mathrm{R}-7.5\mathrm{B}+1}$	842 nm, 668 nm, 475 nm	(Huete et al., 2002)
Chlorophyll Red Edge Index (CIRE)	$\mathrm{CIRE} = \frac{\mathrm{NIR}}{\mathrm{RE}} - 1$	842 nm, 717 nm	(Gitelson et al., 2006)
Simple Ratio (SR)	$\mathrm{SR} = \frac{\mathrm{R}}{\mathrm{NIR}}$	668 nm, 842 nm	(Rouse et al.,1973)
Green Normalized Difference Vegetation Index (GNDVI)	$\mathrm{GNDVI} = \frac{NIR-G}{NIR+G}$	842 nm, 560 nm	(Gitelson et al., 1996)
Chlorophyll Green Index (Clgreen)	$\mathrm{Clgreen} = \frac{\mathrm{NIR}}{\mathrm{G}} - 1$	842 nm, 568 nm	(Gitelson et al., 2003)
Difference Vegetation Index (DVI)	$\mathrm{DVI} = \mathrm{NIR}-\mathrm{R}$	842 nm, 668 nm	(Bannari et al., 1995)
Green optimal Soil Adjusted Vegetation Index (GOSAVI)	$\mathrm{GOSAVI} = \frac{\mathrm{NIR}-\mathrm{G}}{\mathrm{NIR}+\mathrm{G}+0.16}$	842 nm, 560 nm	(Cao et al., 2013)
Optimized Soil Adjusted Vegetation Index (OSAVI)	$\mathrm{OSAVI} = \frac{1.16(\mathrm{NIR}-\mathrm{R})}{\mathrm{NIR}+\mathrm{R}+0.16}$	842 nm, 668 nm	(Rondeaux et al., 1996)
Red -Edge Vegetation Index (RERVI)	$\mathrm{RERVI} = \frac{\mathrm{NIR}}{\mathrm{R}}$	842 nm, 668 nm	(Jordan, 1969)
Ratio Vegetation Index (RVI)	$\mathrm{RVI} = \frac{\mathrm{NIR}}{\mathrm{R}}$	842 nm, 668 nm	(Mishra and Mishra, 2012)
Soil Adjusted Vegetation Index (SAVI)	$\mathrm{SAVI} = 1.5\frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+\mathrm{R}+0.5}$	842 nm, 668 nm	(Huete, 1988)
Structure Insensitive Pigment Index (SIPI)	$\mathrm{SIPI} = \frac{\mathrm{NIR}-\mathrm{B}}{\mathrm{NIR}+\mathrm{R}}$	842 nm, 475 nm	(Penuelas et al., 2013)
Modified Chlorophyll Absorption Reflectance Index (MCARI)	$\mathrm{MCARI} = \mathrm{RE}-\mathrm{R}-0.2(\mathrm{RE}-\mathrm{G})*\frac{\mathrm{RE}}{\mathrm{R}}$	717 nm, 668 nm	(Haboudane et al., 2002b)
Red-edge Normalized Difference vegetation Index (NDRE)	$\mathrm{NDRE} = \frac{\mathrm{NIR}-\mathrm{RE}}{\mathrm{NIR}+\mathrm{RE}}$	842 nm, 717 nm	(Peng and Gitelson, 2012)
Normalized Pigment Chlorophyll Index (NPCI)	$\mathrm{NPCI} = \frac{\mathrm{R}-\mathrm{B}}{\mathrm{R}+\mathrm{B}}$	668 nm, 475 nm	(Merzlyak et al., 1999)
Transformed Chlorophyll Absorption in Reflectance and Soil Adjusted Vegetation Index (TCARI-OSAVI)	$TCARI-OSAVI = \frac{3\left(\mathrm{RE}-\mathrm{R}-0.2(\mathrm{RE}-\mathrm{G})\frac{\mathrm{RE}}{\mathrm{R}}\right)}{(1+0.16)*\frac{\mathrm{NIR}-\mathrm{R}}{\mathrm{NIR}+\mathrm{R}+0.16}}$	717 nm, 668 nm, 842 nm, 560 nm	(Haboudane et al., 2002b)
Normalized Redness Index (NRI)	$\mathrm{NRI} = \frac{\mathrm{R}}{\mathrm{R}+\mathrm{G}+\mathrm{B}}$	668 nm, 560 nm, 475 nm	(Kawashima, 1998)
Green Ratio Vegetation Index (GRVI)	$\mathrm{GRVI} = \frac{\mathrm{NIR}}{\mathrm{G}}$	842 nm, 560 nm	(Wu, 2014)
Plant Sense Reflectance Index (PSRI)	$\mathrm{PSRI} = \frac{\mathrm{R}-\mathrm{G}}{\mathrm{RE}}$	668 nm, 560 nm, 717 nm	(Sims and Gamon, 2002)
Chlorophyll Index (CI)	$\mathrm{CI} = \frac{\mathrm{NIR}}{\mathrm{RE}} - 1$	842 nm, 717 nm	(Abdollahnejad and Panagiotidis, 2020)
Normalized Difference Green/Red Normalized green red difference index (NGRE)	$NGRE = \frac{G-R}{G+R}$	668 nm, 560 nm	(Hunt et al., 2013)
Simple Ratio Pigment Index (SRPI)	$\mathrm{SRPI} = \frac{\mathrm{B}}{\mathrm{R}}$	668 nm, 475 nm	(Cheng et al., 2022)
*Photochemical reflectance index (PRI)	$\mathrm{PRI} = \frac{\mathrm{G}-\mathrm{G}}{\mathrm{G}+\mathrm{G}}$	531 nm, 560 nm	(Penuelas et al., 2013)
*MERIS Terrestrial Chlorophyll Index (MTCI)	$\mathrm{MTCI} = \frac{\mathrm{RE740}-\mathrm{RE705}}{\mathrm{RE}-\mathrm{R}}$	740 nm,705 nm,668 nm	(Dash and Curran, 2004)
*Red edge position (REP)	$\mathrm{REP} = \frac{\mathrm{RE}+40\frac{\mathrm{R}+\mathrm{RE}}{2}-\mathrm{RE}}{\mathrm{RE}-\mathrm{RE}}$	740 nm, 668 nm	(Darvishzadeh et al., 2009)
*Normalized Red-edge index 740–700 (NDRE)	$\mathrm{NDRE} = \frac{\mathrm{RE}-\mathrm{RE}}{\mathrm{RE}+\mathrm{RE}}$	740 nm, 700 nm	(Barnes et al., 2000)

B, R, G, NIR, RE stand for the blue, red, green, near infrared and red edge spectra, respectively. *: Spectral indices calculated only with the ten band sensor.

Table A2
Grey level co-occurrence matrix.

Parameter	Formula	Description
Homogeneity (Hom)	$\sum_{ij=0}^{n-1} \frac{p_{ij}}{1+(i-j)^2}$	Measures the smoothness of image texture (Tuttle et al., 2006)
Correlation (Corr)	$\sum_{ij=0}^{n-1} P_{ij}\left(\frac{(i-\mu_i)(i-\mu_i)}{(\partial_i^2)(\partial_i^2)}\right)$	Measures the grey level linear dependency within an image (Kayitakire et al., 2006)
Dissimilarity (Diss)	$\sum_{ij=0}^{M-1} P_{ij}\lvert i-j\rvert$	Measure of local variation (Rubner et al., 2001)
Variance (Var)	$\sigma_i^2 = \sum_{ij=0}^{n-1} P_{ij}(i-\mu)^2$	Variability of spectral response of pixels (Materka and Strzelecki, 1998)
Entropy (Ento)	$\sum_{ij=0}^{n-1} [\![P_{ij}(-]\!] lnP_{ij})$	Statistical measure of uncertainty (Yuan et al., 1991)
Mean (Mean)	$\mu_i = \sum_{ij=0}^{n-1} i(P_{ij})$ $\mu_j = \sum_{ij=0}^{n-1} i(P_{ij})$	Average grey-level in the minor neighbourhood (Materka and Strzelecki, 1998)
Dissimilarity (Diss)	$\sum_{ij=0}^{n-1} P_{ij}$	Measures local variation (Rubner et al., 2001)
Second moment (Sec)	$\sum_{ij=0}^{n-1} P_{ij^2}$	Indicator of local homogeneity (Lottering et al., 2019).

Where P(i,j) is the normalized co-occurrence matrix where the sum of (i,j = 0, n-1)(P(i,j)) = 1.

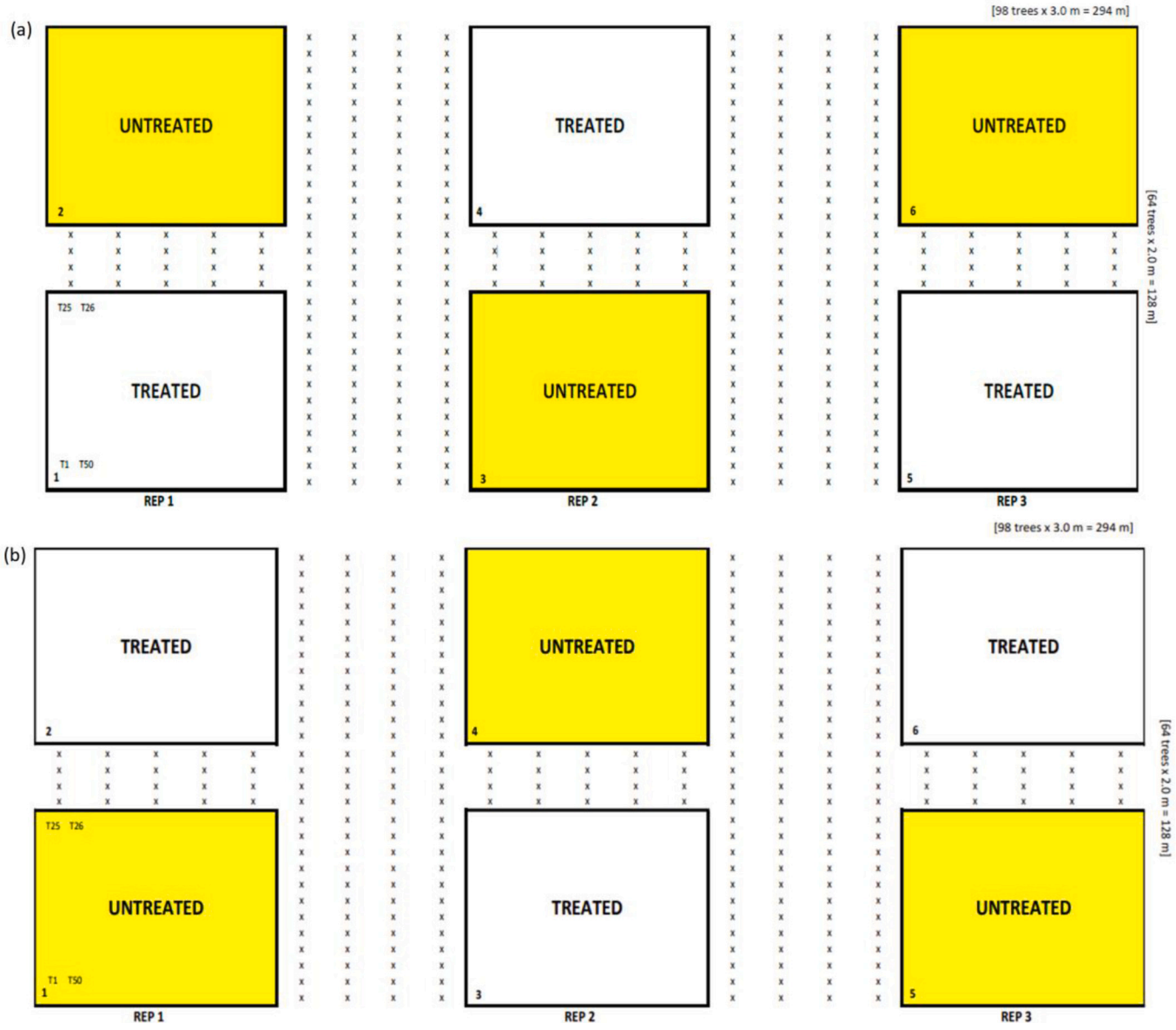


Fig. A1. *Gonipterus* sp. n. 2 pest exclusion experimental trial layout and setup (a) Greytown, (b) Ixopo.

Table A3
Statistical analysis of vegetation indices using ANOVA.

Indices	Statistical test	Post hoc test *	Damage	Site	Damage x Site	No damage	Low	Medium	High
CI	ANOVA	T	0.98	0.85	0.85				
EVI	KW	D	**<0.001**			**b**	**b**	**a**	**a**
GNDVI	KW	D	**<0.001**			a	b	c	d
MCARI	KW	D	**<0.001**			**a**	**ab**	**b**	**a**
NDRE740	ANOVA	T	**<0.001**	**<0.001**	**<0.001**	**c**	**b**	**a**	**a**
NDVI	KW	D	**<0.001**	**<0.001**	**<0.001**	**a**	**b**	**c**	**d**
Tcari_Osavi	KW	D	**<0.001**						
ARI	ANOVA	T	0.90	0.90	<0.55				
NGRE	ANOVA	D	**<0.001**	**<0.001**	**<0.001**	**a**	**b**	**c**	**c**
Clgreen	KW	D	**<0.001**						
RVI	KW	D	**<0.001**			d	c	b	a
OSAVI	ANOVA	T	0.16	0.06	**<0.001**				
DVI	KW	D	**<0.001**			d	c	b	a
CIRE	KW	D	**<0.001**			a	b	a	a
GOSAVI	KW	D	0.164						
SAVI	KW	D	**<0.001**			**b**	**b**	**a**	**a**
SR	KW	D	**<0.001**			**b**	**c**	**a**	**d**
LCI	KW	D	**<0.001**			b	a	c	a
SRPI	KW	D	**<0.001**			b	b	a	a
NRI	KW	D	**<0.001**			a	b	c	c
GRVI	KW	D	**<0.001**						
PSRI	KW	D	**<0.001**			a	a	a	b
NPCI	KW	D	**<0.001**			c	b	d	a
RERVI	KW	D	**<0.001**			b	a	a	a
PRI	KW	D	**<0.001**			a	a	b	a
REP	ANOVA	T	0.44	0.74	**<0.05**				
RE750–700	KW	D	**<0.001**			a	b	c	c
MTCI	ANOVA	T	0.76	0.76	0.98				
SIPI	ANOVA	T	**<0.001**	**<0.001**	**<0.001**	**c**	**b**	**d**	**a**

* T = Tukey, D = Duncan.

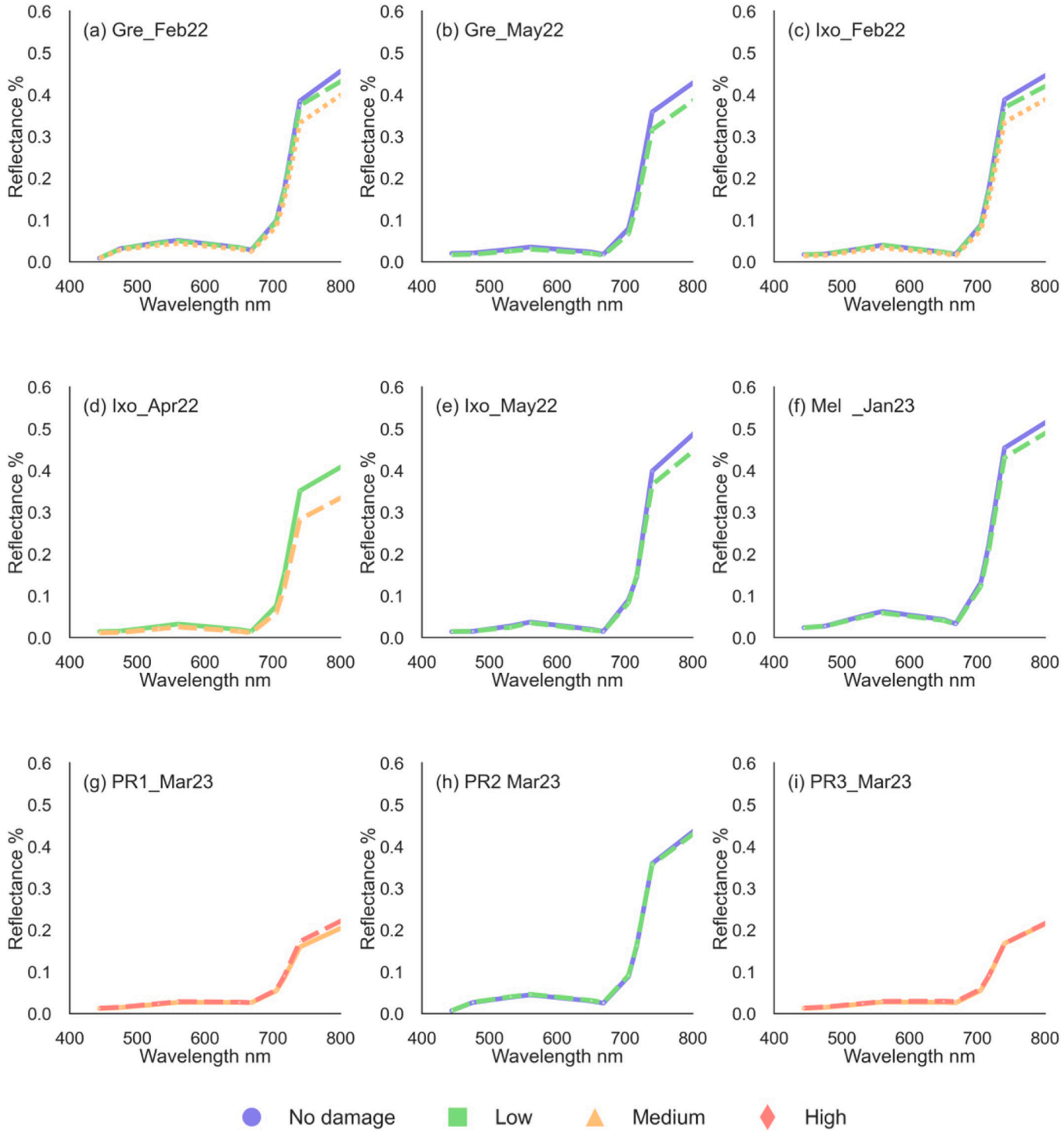


Fig. A2. Average Spectral reflectance across multiple sites based on five and ten band cameras and the level of *Gonipterus* sp. n. 2 damage before histogram matching.

Table A4
Overall accuracies for each model for the five and 10-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) before histogram matching using XGBoost classifier.

Model	Accuracy	F1-score	Precision	Recall
10 Bands				
Spectra	79 %	0.79	0.79	0.79
VIs	77 %	0.77	0.77	0.77
Texture	65 %	0.65	0.65	0.65
VIs + Spectra	84 %	0.84	0.84	0.84
VIs + Spectra + Texture	86 %	0.86	0.86	0.86
5 Bands				
Spectra	75 %	0.75	0.75	0.75
VIs	76 %	0.76	0.76	0.76
Texture	61 %	0.61	0.62	0.62
VIs + Spectra	79 %	0.79	0.79	0.79
VIs + Spectra + Texture	83 %	0.83	0.83	0.83

Table A5
Overall accuracies for each model for the five and 10-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) before histogram matching using RF Classification.

Model	Accuracy	F1-score	Precision	Recall
10 Bands				
Spectra	80%	0.80	0.80	0.80
VIs	79 %	0.79	0.79	0.79
Texture	67 %	0.67	0.67	0.67
VIs + Spectra	83 %	0. 83	0. 83	0. 83
VIs + Spectra + Texture	89 %	0.89	0.89	0.89
5 Bands				
Spectra	77 %	0.77	0.77	0.77
VIs	77 %	0.77	0.77	0.77
Texture	66 %	0.66	0.66	0.66
VIs + Spectra	79 %	0.79	0.79	0.79
VIs + Spectra + Texture	83 %	0.83	0.83	0.83

Table A6
Overall accuracies for each model for the 10 and 5-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) after histogram matching using RF Classification.

Model	Accuracy	F1-score	Precision	Recall
10 Bands				
Spectra	84 %	0.84	0.84	0.84
VIs	81 %	0.81	0.81	0.81
Texture	76 %	0.76	0.76	0.76
VIs + Spectra	83 %	0.83	0.83	0.83
VIs + Spectra + Texture	90 %	0.90	0.90	0.90
Reduced 5 bands				
Spectra	81 %	0.81	0.81	0.81
VIs	80%	0.81	0.81	0.81
Texture	75 %	0.75	0.75	0.75
VIs + Spectra	81 %	0.81	0.81	0.81
VIs + Spectra + Texture	89 %	0.89	0.89	0.89

Table A7
Overall accuracies for each model for the five and 10-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) before histogram matching using SVM.

Model	Accuracy	F1-score	Precision	Recall
10 Bands				
Spectra	62 %	0.63	0.64	0.65
VIs	63 %	0.63	0.63	0.63
Texture	66 %	0.65	0.66	0.67

(*continued on next page*)

Table A7 (*continued*)

Model	Accuracy	F1-score	Precision	Recall
Vis + Spectra	69 %	0.68	0.68	0.69
VIs + Spectra + Texture	77 %	0.77	0.77	0.77
Reduced 5 bands				
Spectra	66 %	0.65	0.65	0.66
VIs	60 %	0.59	0.58	0.60
Texture	64 %	0.63	0.64	0.65
Vis + Spectra	64 %	0.63	0.63	0.64
VIs + Spectra + Texture	71 %	0.71	0.71	0.71

Table A8
Overall accuracies for each model for the five and 10-band independent test datasets based on 10-band data (Micasense-MX RedEdge Dual), 5-band data (same MicaSense-MX RedEdge Dual datasets, only the 5 channels of the 'red' band were selected) after histogram matching using SVM.

Model	Accuracy	F1-score	Precision	Recall
10 Bands				
Spectra	66 %	0.66	0.62	0.66
VIs	78 %	0.77	0.77	0.78
Texture	73 %	0.72	0.72	0.73
Vis + Spectra	79 %	0.79	0.79	0.79
VIs + Spectra + Texture	85 %	0.85	0.85	0.85
5 bands				
Spectra	64 %	0.57	0.56	0.64
VIs	71 %	0.69	0.70	0.71
Texture	70 %	0.69	0.70	0.70
VIs + Spectra	72 %	0.71	0.72	0.72
VIs + Spectra + Texture	78 %	0.78	0.78	0.78

Table A9
Overall leave one site out accuracies for the full 10 band data from MicaSense-MX RedEdge Dual datasets using RF Classification.

Test dataset Site	Accuracy	Precision weighted	Recall weighted	F1 weighted
Gre_Feb22	39 %	0.26	0.39	0.26
Gre_May22	37 %	0.62	0.37	0.46
Ixo_Feb22	62 %	0.64	0.62	0.58
Ixo_Apr22	26 %	0.48	0.26	0.16
Ixo_May22	51 %	0.79	0.51	0.56
Mel_Jan23	35 %	0.53	0.35	0.21
PR1_Mar23	9 %	0.37	0.09	0.04
PR2_Mar23	27 %	0.69	0.27	0.34
PR3_Mar23	16 %	0.93	0.16	0.26

Table A10
Overall leave one site out accuracies for the full 10 band data from MicaSense-MX RedEdge Dual datasets using SVM.

Test dataset Site	Accuracy	Precision weighted	Recall weighted	F1 weighted
Gre_Feb22	51 %	0.71	0.51	0.47
Gre_May22	34 %	0.63	0.34	0.43
Ixo_Feb22	48 %	0.73	0.48	0.48
Ixo_Apr22	31 %	0.74	0.31	0.28
Ixo_May22	61 %	0.76	0.61	0.66
Mel_Jan23	9 %	0.55	0.09	0.13
PR1_Mar23	3 %	0.55	0.03	0.02
PR2_Mar23	39 %	0.81	0.39	0.40
PR3_Mar 23	26 %	0.93	0.26	0.41

Data availability

The scripts used in this study are publicly available at https://github.com/Pollen1/Assessing-Gonipterus-damage-using-UAV-data/tree/main. The datasets used are subject to confidentiality agreements with forestry companies and research partners. However, the data can be made available by the corresponding author and with permission from the relevant stakeholders.

References

Abdollahnejad, A., Panagiotidis, D., 2020. Tree species classification and health status assessment for a mixed broadleaf-conifer Forest with UAS multispectral imaging. Remote Sens. 12, 3722. https://doi.org/10.3390/rs12223722.

Adelabu, S., Mutanga, O., Adam, E., 2014. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. ISPRS J. Photogramm. Remote Sens. 95, 34–41. https://doi.org/10.1016/j.isprsjprs.2014.05.013.

Bagnall, G.C., Thomasson, J.A., Yang, C., Wang, T., Han, X., Sima, C., Chang, A., 2023. Uncrewed aerial vehicle radiometric calibration: a comparison of autoexposure and fixed-exposure images. Plant Phenome J. 6, e20082. https://doi.org/10.1002/ppj2.20082.

Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. Remote Sens. Rev. 13, 95–120. https://doi.org/10.1080/02757259509532298.

Barnes, E., Clarke, T.R., Richards, S.E., Colaizzi, P., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T.L., 2000. Coincident Detection of Crop Water Stress, Nitrogen Status, and Canopy Density Using Ground Based Multispectral Data.

Borra-Serrano, I., Peña, J.M., Torres-Sánchez, J., Mesas-Carrascosa, F.J., López-Granados, F., 2015. Spatial quality evaluation of resampled unmanned aerial vehicle-imagery for weed mapping. Sensors 15, 19688–19708. https://doi.org/10.3390/s150819688.

Bradshaw, C.J.A., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., Barbet-Massin, M., Salles, J.-M., Simard, F., Courchamp, F., 2016. Massive yet grossly underestimated global costs of invasive insects. Nat. Commun. 7, 12986. https://doi.org/10.1038/ncomms12986.

Brockerhoff, E.G., Corley, J.C., Jactel, H., Miller, D.R., Rabaglia, R.J., Sweeney, J., Paine, T.D., Slippers, B., Wingfield, M.J., 2023. Monitoring and surveillance of forest insects. In: Allison, D. (Ed.), Forest Entomology and Pathology, Vol. 1. Entomology. Springer International Publishing, Cham, pp. 669–705. https://doi.org/10.1007/978-3-031-11553-0_19.

Canelles, Q., Aquilué, N., James, P.M.A., Lawler, J., Brotons, L., 2021. Global review on interactions between insect pests and other forest disturbances. Landsc. Ecol. 36, 945–972. https://doi.org/10.1007/s10980-021-01209-7.

Cao, Q., Miao, Y., Wang, H., Huang, S., Cheng, S., Khosla, R., Jiang, R., 2013. Non-destructive estimation of rice plant nitrogen status with crop circle multispectral active canopy sensor. Field Crop Res. 154, 133–144. https://doi.org/10.1016/j.fcr.2013.08.005.

Chabalala, Y., Adam, E., Ali, K.A., 2023. Exploring the effect of balanced and imbalanced multi-class distribution data and sampling techniques on fruit-tree crop classification using different machine learning classifiers. Geomatics 3, 70–92. https://doi.org/10.3390/geomatics3010004.

Chambi-Legoas, R., Tomazello-Filho, M., Guedes, F.T.P., Chaix, G., 2022. High growth recovery ability of *Eucalyptus grandis* trees following a 3-year period of 80% throughfall reduction. For. Ecol. Manag. 503, 119766. https://doi.org/10.1016/j.foreco.2021.119766.

Cheng, M., Jiao, X., Liu, Y., Shao, M., Yu, X., Bai, Y., Wang, Z., Wang, S., Tuohuti, N., Liu, S., Shi, L., Yin, D., Huang, X., Nie, C., Jin, X., 2022. Estimation of soil moisture content under high maize canopy coverage from UAV multimodal data and machine learning. Agric. Water Manag. 264, 107530. https://doi.org/10.1016/j.agwat.2022.107530.

Conradie, D.C.U., 2012. South Africa's Climatic Zones: Today, Tomorrow. South Africa.

Dainelli, R., Toscano, P., Di Gennaro, S.F., Matese, A., 2021. Recent advances in unmanned aerial vehicles forest remote sensing—a systematic review. Part II: Research applications. Forests 12, 397. https://doi.org/10.3390/f12040397.

Daniels, L., Eeckhout, E., Wieme, J., Dejaegher, Y., Audenaert, K., Maes, W.H., 2023. Identifying the optimal radiometric calibration method for UAV-based multispectral imaging. Remote Sens. 15, 2909. https://doi.org/10.3390/rs15112909.

Darvishzadeh, R., Atzberger, C., Skidmore, A.K., Abkar, A.A., 2009. Leaf area index derivation from hyperspectral vegetation indicesand the red edge position. Int. J. Remote Sens. 30, 6199–6218. https://doi.org/10.1080/01431160902842342.

Dash, J., Curran, P.J., 2004. The MERIS terrestrial chlorophyll index. Int. J. Remote Sens. 25, 5403–5413. https://doi.org/10.1080/0143116042000274015.

Datt, B., 1999. A new reflectance index for remote sensing of chlorophyll content in higher plants: tests using Eucalyptus leaves. J. Plant Physiol. 154, 30–36. https://doi.org/10.1016/S0176-1617(99)80314-9.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., de Colstoun, E.B., McMurtrey, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74, 229–239. https://doi.org/10.1016/S0034-4257(00)00113-9.

Debeurs, K., Townsend, P., 2008. Estimating the effect of gypsy moth defoliation using MODIS. Remote Sens. Environ. 112, 3983–3990. https://doi.org/10.1016/j.rse.2008.07.008.

Devi, D., Biswas, S.K., Purkayastha, B., 2019. Learning in presence of class imbalance and class overlapping by using one-class SVM and undersampling technique. Connect. Sci. 31, 105–142. https://doi.org/10.1080/09540091.2018.1560394.

dos Santos, A., Oumar, Z., Arnhold, A., da Silva, N., Oliveira Silva, C., Zanetti, R., 2017. Multispectral characterization, prediction and mapping of *Thaumastocoris peregrinus* (Hemiptera: Thamascoridae) attack in Eucalyptus plantations using remote sensing. J. Spat. Sci. 62, 127–137. https://doi.org/10.1080/14498596.2016.1220873.

Estrada, J.S., Fuentes, A., Reszka, P., Auat Cheein, F., 2023. Machine learning assisted remote forestry health assessment: a comprehensive state of the art review. Front. Plant Sci. 14.

Garcia, A., Samalens, J.-C., Grillet, A., Soares, P., Branco, M., Van Halder, I., Jactel, H., Battisti, A., 2023. Testing early detection of pine processionary moth Thaumetopoea pityocampa nests using UAV-based methods. NB 84, 267–279. https://doi.org/10.3897/neobiota.84.95692.

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens. Environ. 58, 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7.

Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys. Res. Lett. 30. https://doi.org/10.1029/2002GL016450.

Gitelson, A.A., Keydan, G.P., Merzlyak, M.N., 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. Geophys. Res. Lett. 33. https://doi.org/10.1029/2006GL026457.

Gitelson, A.A., Chivkunova, O.B., Merzlyak, M.N., 2009. Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. Am. J. Bot. 96, 1861–1868. https://doi.org/10.3732/ajb.0800395.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 81, 416–426. https://doi.org/10.1016/S0034-4257(02)00018-4.

Health (PLH), E.P. on P, Jeger, M., Bragard, C., Caffier, D., Candresse, T., Chatzivassiliou, E., Dehnen-Schmutz, K., Gilioli, G., Jaques Miret, J.A., MacLeod, A., Navajas Navarro, M., Niere, B., Parnell, S., Potting, R., Rafoss, T., Rossi, V., Urek, G., Van Bruggen, A., Van der Werf, W., West, J., Winter, S., Santolamazza-Carbone, S., Kertész, V., Aukhojee, M., Grégoire, J.-C., 2018. Pest categorisation of the Gonipterus scutellatus species complex. EFSA J. 16, e05107. https://doi.org/10.2903/j.efsa.2018.5107.

Heim, R.H.J., Okole, N., Steppe, K., Van Labeke, M.-C., Geedicke, I., Maes, W.H., 2024. An applied framework to unlocking multi-angular UAV reflectance data: a case study for classification of plant parameters in maize (*Zea mays*). Precis. Agric. 25, 1751–1775. https://doi.org/10.1007/s11119-024-10133-0.

Helmer, E.H., Ruefenacht, B., 2005. Cloud-free satellite image mosaics with regression trees and histogram matching. Photogramm. Eng. Remote. Sens. 71, 1079–1089. https://doi.org/10.14358/PERS.71.9.1079.

Hennessy, A., Clarke, K., Lewis, M., 2020. Hyperspectral classification of plants: a review of waveband selection generalisability. Remote Sens. 12, 113. https://doi.org/10.3390/rs12010113.

Huang, L., Liu, Y., Huang, W., Dong, Y., Ma, H., Wu, K., Guo, A., 2022. Combining random Forest and XGBoost methods in detecting early and mid-term winter wheat stripe rust using canopy level hyperspectral measurements. Agriculture 12, 74. https://doi.org/10.3390/agriculture12010074.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

Hunt, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M., Akhmedov, B., 2013. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. Int. J. Appl. Earth Obs. Geoinf. 21, 103–112. https://doi.org/10.1016/j.jag.2012.07.020.

Ismail, R., Mutanga, O., 2010. A comparison of regression tree ensembles: predicting Sirex noctilio induced water stress in *Pinus patula* forests of KwaZulu-Natal, South Africa. Int. J. Appl. Earth Obs. Geoinf. 12, S45–S51. https://doi.org/10.1016/j.jag.2009.09.004.

Jafarbiglu, H., Pourreza, A., 2023. Impact of sun-view geometry on canopy spectral reflectance variability. ISPRS J. Photogramm. Remote Sens. 196, 270–286. https://doi.org/10.1016/j.isprsjprs.2022.12.002.

Jordan, C.F., 1969. Derivation of leaf-area index from quality of light on the Forest floor. Ecology 50, 663–666. https://doi.org/10.2307/1936256.

Joubert, J., Sivparsad, B., Schröder, M., Germishuizen, I., Chen, J., Hurley, B., Allison, J. D., Hammerbacher, A., 2023. Susceptibility of Eucalyptus trees to defoliation by the Eucalyptus snout beetle, Gonipterus sp. n. 2, is enhanced by high foliar contents of 1,8-cineole, oxalic acid and sucrose and low contents of palmitic and shikimic acid. Plant Cell Environ. 46, 3481–3500. https://doi.org/10.1111/pce.14696.

Kasinathan, T., Uyyala, S.R., 2021. Machine learning ensemble with image processing for pest identification and classification in field crops. Neural Comput. & Applic. 33, 7491–7504. https://doi.org/10.1007/s00521-020-05497-z.

Kawashima, S., 1998. An algorithm for estimating chlorophyll content in leaves using a video camera. Ann. Bot. 81, 49–54. https://doi.org/10.1006/anbo.1997.0544.

Kayitakire, F., Hamel, C., Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. Remote Sens. Environ. 102, 390–401. https://doi.org/10.1016/j.rse.2006.02.022.

Keetse, H., 2024. Tree Growth Losses from Gonipterus Sp. 2 Feeding Damage on Eucalyptus Plantations and Quantified Economic Impact. University of Pretoria, Pretoria.

Kelsey, K.C., Neff, J.C., 2014. Estimates of aboveground biomass from texture analysis of Landsat imagery. Remote Sens. 6, 6407–6422. https://doi.org/10.3390/rs6076407.

Kimura, R., Okada, S., Miura, H., Kamichika, M., 2004. Relationships among the leaf area index, moisture availability, and spectral reflectance in an upland rice field. Agric. Water Manag. 69, 83–100. https://doi.org/10.1016/j.agwat.2004.04.009.

Kume, A., 2017. Importance of the green color, absorption gradient, and spectral absorption of chloroplasts for the radiative energy balance of leaves. J. Plant Res. 130, 501–514. https://doi.org/10.1007/s10265-017-0910-z.

Lausch, A., Erasmi, S., King, D., Magdon, P., Heurich, M., 2016. Understanding forest health with remote sensing -Part I—A review of spectral traits. In: Processes and Remote-Sensing Characteristics. Remote Sensing, 8, p. 1029. https://doi.org/10.3390/rs8121029.

Lauwers, M., De Cauwer, B., Nuyttens, D., Maes, W.H., Pieters, J.G., 2024. Multispectral UAV image classification of jimson weed (*Datura stramonium* L.) in common bean (*Phaseolus vulgaris* L.). Remote Sens. 16, 3538. https://doi.org/10.3390/rs16183538.

Liao, K., Yang, F., Dang, H., Wu, Y., Luo, K., Li, G., 2022. Detection of Eucalyptus leaf disease with UAV multispectral imagery. Forests 13, 1322. https://doi.org/10.3390/f13081322.

Lin, Q., Huang, H., Wang, J., Huang, K., Liu, Y., 2019. Detection of pine shoot beetle (PSB) stress on pine forests at individual tree level using UAV-based hyperspectral imagery and Lidar. Remote Sens. 11, 2540. https://doi.org/10.3390/rs11212540.

Lin, Q., Huang, H., Wang, J., Chen, L., Du, H., Zhou, G., 2023. Early detection of pine shoot beetle attack using vertical profile of plant traits through UAV-based hyperspectral, thermal, and lidar data fusion. Int. J. Appl. Earth Obs. Geoinf. 125, 103549. https://doi.org/10.1016/j.jag.2023.103549.

Liu, T., Shi, T., Zhang, H., Wu, C., 2020. Detection of rise damage by leaf folder (Cnaphalocrocis medinalis) using unmanned aerial vehicle based hyperspectral data. Sustainability 12, 9343. https://doi.org/10.3390/su12229343.

Lottering, R., Mutanga, O., 2016. Optimising the spatial resolution of WorldView-2 pan-sharpened imagery for predicting levels of Gonipterus scutellatus defoliation in KwaZulu-Natal, South Africa. ISPRS J. Photogramm. Remote Sens. 112, 13–22. https://doi.org/10.1016/j.isprsjprs.2015.11.010.

Lottering, R., Mutanga, O., Peerbhay, K., Ismail, R., 2019. Detecting and mapping Gonipterus scutellatus induced vegetation defoliation using WorldView-2 pan-sharpened image texture combinations and an artificial neural network. JARS 13, 014513. https://doi.org/10.1117/1.JRS.13.014513.

Lottering, R., Mutanga, O., Peerbhay, K., Lottering, S., 2020. Spatially optimizing vegetation indices integrated with sparse partial least squares regression to detect and map the effects of *Gonipterus scutellatus* on the chlorophyll content of eucalyptus plantations. Int. J. Remote Sens. 41, 6444–6459. https://doi.org/10.1080/01431161.2020.1739350.

Luo, Y., Huang, H., Roques, A., 2023. Early monitoring of Forest wood-boring pests with remote sensing. Annu. Rev. Entomol. 68, 277–298. https://doi.org/10.1146/annurev-ento-120220-125410.

Ma, Y., Lu, J., Huang, X., 2023. Damage diagnosis of Pinus yunnanensis canopies attacked by Tomicus using UAV hyperspectral images. Forests 14, 61. https://doi.org/10.3390/f14010061.

Maes, W.H., 2025. Practical guidelines for performing UAV mapping flights with snapshot sensors. Remote Sens. 17, 606. https://doi.org/10.3390/rs17040606.

Mahapatra, D., 2014. Analyzing training information from random forests for improved image segmentation. IEEE Trans. Image Process. 23, 1504–1512. https://doi.org/10.1109/TIP.2014.2305073.

Mally, C.W., 1924. The eucalyptus snout-beetle. J. Departm. Agric. 9, 415–442. https://doi.org/10.10520/AJA0000020_683.

Mapondera, T.S., Burgess, T., Matsuki, M., Oberprieler, R.G., 2012. Identification and molecular phylogenetics of the cryptic species of the *Gonipterus scutellatus* complex (Coleoptera: Curculionidae: Gonipterini): resolving the *Gonipterus scutellatus* complex. Aust. J. Entomol. 51, 175–188. https://doi.org/10.1111/j.1440-6055.2011.00853.x.

Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Rakitin, V.Y.U., 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. Physiol. Plant. 106, 135–141. https://doi.org/10.1034/j.1399-3054.1999.106119.x.

Mielke, M.S., Schaffer, B., Schilling, A.C., 2012. Evaluation of reflectance spectroscopy indices for estimation of chlorophyll content in leaves of a tropical tree species. Photosynthetica 50, 343–352. https://doi.org/10.1007/s11099-012-0038-2.

Minařík, R., Langhammer, J., Lendzioch, T., 2021. Detection of bark beetle disturbance at tree level using UAS multispectral imagery and deep learning. Remote Sens. 13, 4768. https://doi.org/10.3390/rs13234768.

Mishra, S., Mishra, D.R., 2012. Normalized difference chlorophyll index: a novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. Remote Sens. Environ. 117, 394–406. https://doi.org/10.1016/j.rse.2011.10.016.

Mngadi, M., Germishuizen, I., Mutanga, O., Naicker, R., Maes, W.H., Odebiri, O., Schroder, M., 2024. A systematic review of the application of remote sensing technologies in mapping forest insect pests and diseases at a tree-level. Remote Sens. Appl. Soc. Environ. 36, 101341. https://doi.org/10.1016/j.rsase.2024.101341.

Moley, L.M., Goodin, D.G., Winslow, W.P., 2022. Leaf-level spectroscopy for analysis of invasive Pest impact on trees in a stressed environment: an example using emerald ash borer (*Agrilus planipennis Fairmaire*) in ash trees (Fraxinus spp.), Kansas, USA. Environments 9, 42. https://doi.org/10.3390/environments9040042.

Morris, A.R., 2022. Changing use of species and hybrids in south African forest plantations. South. For. J. For. Sci. 84, 193–205. https://doi.org/10.2989/20702620.2022.2110538.

Newete, S.W., Oberprieler, R.G., Byrne, M.J., 2011. The host range of the Eucalyptus Weevil, Gonipterus "scutellatus" Gyllenhal (Coleoptera: Curculionidae), in South Africa. Ann. For. Sci. 68, 1005–1013. https://doi.org/10.1007/s13595-011-0108-9.

Nguyen, C., Sagan, V., Skobalski, J., Severo, J.I., 2023. Early detection of wheat yellow rust disease and its impact on terminal yield with multi-spectral UAV-imagery. Remote Sens. 15, 3301. https://doi.org/10.3390/rs15133301.

Pan, J., Lin, J., Xie, T., 2023. Exploring the potential of UAV-based hyperspectral imagery on pine wilt disease detection: influence of Spatio-temporal scales. Remote Sens. 15, 2281. https://doi.org/10.3390/rs15092281.

Panopoulou, C., Antonopoulos, A., Arapostathi, E., Stamouli, M., Katsileros, A., Tsagkarakis, A., 2024. Using multispectral data from UAS in machine learning to detect infestation by Xylotrechus chinensis (Chevrolat) (Coleoptera: Cerambycidae) in mulberries. Agronomy 14, 2061. https://doi.org/10.3390/agronomy14092061.

Peng, Y., Gitelson, A.A., 2012. Remote estimation of gross primary productivity in soybean and maize based on total crop chlorophyll content. In: Remote Sensing of Environment, Remote Sensing of Urban Environments, 117, pp. 440–448. https://doi.org/10.1016/j.rse.2011.10.021.

Penuelas, J., Frederic, B., Filella, I., 2013. Photosyn1995.

Quan, Y., Zhong, X., Feng, W., Chan, J.C.-W., Li, Q., Xing, M., 2021. SMOTE-based weighted deep rotation Forest for the imbalanced hyperspectral data classification. Remote Sens. 13, 464. https://doi.org/10.3390/rs13030464.

Reis, A.R., Ferreira, L., Tomé, M., Araujo, C., Branco, M., 2012. Efficiency of biological control of *Gonipterus platensis* (Coleoptera: Curculionidae) by *Anaphes nitens* (Hymenoptera: Mymaridae) in cold areas of the Iberian Peninsula: implications for defoliation and wood production in *Eucalyptus globulus*. For. Ecol. Manag. 270, 216–222. https://doi.org/10.1016/j.foreco.2012.01.038.

Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55, 95–107. https://doi.org/10.1016/0034-4257(95)00186-7.

Rubner, Y., Puzicha, J., Tomasi, C., Buhmann, J.M., 2001. Empirical evaluation of dissimilarity measures for color and texture. Comput. Vis. Image Underst. 84, 25–43. https://doi.org/10.1006/cviu.2001.0934.

Schröder, M.L., Nahrung, H.F., Lawson, S.A., Slippers, B., Wingfield, M.J., Hurley, B.P., 2017. Biological control of the *Gonipterus scutellatus* species complex: Testing the species, climatic or phenological mismatch hypotheses. In: Mason, P.G., Gillespie, D. R., Vincent, C. (Eds.), Proceedings of the 5th International Symposium on Biological Control of Arthropods, Langkawi, Malaysia, September 11–15, 2017. CABI, UK, pp. 99–101. https://doi.org/10.1079/9781786394118.0099.

Schröder, M.L., Slippers, B., Wingfield, M.J., Hurley, B.P., 2020. Invasion history and management of Eucalyptus snout beetles in the Gonipterus scutellatus species complex. J. Pest. Sci. 93, 11–25. https://doi.org/10.1007/s10340-019-01156-y.

Senf, C., Seidl, R., Hostert, P., 2017. Remote sensing of forest insect disturbances: current state and future directions. Int. J. Appl. Earth Obs. Geoinf. 60, 49–60. https://doi.org/10.1016/j.jag.2017.04.004.

Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81, 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X.

Stone, C., Mohammed, C., 2017. Application of remote sensing technologies for assessing planted forests damaged by insect pests and fungal pathogens: a review. Curr. For. Rep. 3, 75–92. https://doi.org/10.1007/s40725-017-0056-1.

Stone, C., Chisholm, L., Coops, N., 2001. Spectral reflectance characteristics of eucalypt foliage damaged by insects. Aust. J. Bot. 49, 687–698. https://doi.org/10.1071/bt00091.

Swaminathan, V., Thomasson, J.A., Hardin, R.G., Rajan, N., Raman, R., 2024. Selection of appropriate multispectral camera exposure settings and radiometric calibration methods for applications in phenotyping and precision agriculture. Plant Phenome J. 7, e70000. https://doi.org/10.1002/ppj2.70000.

Tooke, F.G.C., 1955. The Eucalyptus Snout. Beetle, Gonipterus scutellatus Gyll. A Study of its Ecology and Control by biological Means. The Eucalyptus Snout.beetle, Gonipterus scutellatus Gyll. A Study of its Ecology and Control by Biological Means.

Tribe, G.D., 2005. The present status of Anaphes nitens (Hymenoptera: Mymaridae), an egg parasitoid of the Eucalyptus snout beetle Gonipterus scutellatus, in the Western Cape Province of South Africa. South. Afr. For. J. 203, 49–54. https://doi.org/10.2989/10295920509505218.

Tuttle, E.M., Jensen, R.R., Formica, V.A., Gonser, R.A., 2006. Using remote sensing image texture to study habitat use patterns: a case study using the polymorphic white-throated sparrow (*Zonotrichia albicollis*). Glob. Ecol. Biogeogr. 15, 349–357. https://doi.org/10.1111/j.1466-822X.2006.00232.x.

Williams, G.M., Ginzel, M.D., Ma, Z., Adams, D.C., Campbell, F., Lovett, G.M., Pildain, M. B., Raffa, K.F., Gandhi, K.J.K., Santini, A., Sniezko, R.A., Wingfield, M.J., Bonello, P., 2023. The global forest health crisis: a public-good social dilemma in need of international collective action. Annu. Rev. Phytopathol. 61, 377–401. https://doi.org/10.1146/annurev-phyto-021722-024626.

Wingfield, M., Slippers, B., Hurley, B., Coutinho, T., Wingfield, B., Roux, J., 2008. Eucalypt pests and diseases: growing threats to plantation productivity. South. For. J. For. Sci. 70, 139–144. https://doi.org/10.2989/SOUTH.FOR.2008.70.2.9.537.

Wu, W., 2014. The generalized difference vegetation index (GDVI) for dryland characterization. Remote Sens. 6, 1211–1233. https://doi.org/10.3390/rs6021211.

Wu, D., Yu, L., Yu, R., Zhou, Q., Li, J., Zhang, X., Ren, L., Luo, Y., 2023. Detection of the monitoring window for pine wilt disease using multi-temporal UAV-based multispectral imagery and machine learning algorithms. Remote Sens. 15, 444. https://doi.org/10.3390/rs15020444.

Xu, H., Wei, Yuchun, Li, Xiao, Zhao, Yadi, Cheng, Q., 2021. A novel automatic method on pseudo-invariant features extraction for enhancing the relative radiometric normalization of high-resolution images. Int. J. Remote Sens. 42, 6153–6183. https://doi.org/10.1080/01431161.2021.1934912.

Xu, Z., Zhang, Q., Xiang, S., Li, Y., Huang, X., Zhang, Y., Zhou, X., Li, Z., Yao, X., Li, Q., Guo, X., 2022. Monitoring the severity of Pantana phyllostachysae Chao infestation in Moso bamboo forests based on UAV multi-spectral remote sensing feature selection. Forests 13, 418. https://doi.org/10.3390/f13030418.

Xulu, S., Mbatha, N., Peerbhay, K., Gebreslasie, M., Agjee, N., 2024. Comparison of different spectral indices to differentiate the impact of insect attack on planted forest stands. Remote Sens. Appl. Soc. Environ. 33, 101087. https://doi.org/10.1016/j.rsase.2023.101087.

Yuan, X., King, D., Vlcek, J., 1991. Sugar maple decline assessment based on spectral and textural analysis of multispectral aerial videography. Remote Sens. Environ. 37, 47–54. https://doi.org/10.1016/0034-4257(91)90049-C.