Remote Sensing in Ecology and Conservation
Open Access
ZSL LET'S WORK FOR WILDLIFE

ORIGINAL RESEARCH

Monitoring ash dieback (*Hymenoscyphus fraxineus*) in British forests using hyperspectral remote sensing

Aland H. Y. Chan[1], Chloe Barnes[2], Tom Swinfield[1] & David A. Coomes[1]

[1]Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK
[2]2Excel geo, Hall Farm 2, Sywell Aerodrome, Sywell NN6 0BN, UK

Keywords
hyperspectral remote sensing, ash dieback, *Hymenoscyphus fraxineus*, temperate forests, dark pixel filtering, species classification

Correspondence
David A. Coomes, Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Downing Street, CB2 3EA, UK. Tel: (+44) 01223 333911; Fax:(+44) 01223 332517
Email: dac18@cam.ac.uk

Editor: Kate He
Associate Editor: András Zlinszky

Received: 27 April 2020; Revised: 9 September 2020; Accepted: 25 November 2020

doi: 10.1002/rse2.190

Remote Sensing in Ecology and Conservation 2021;**7** (2):306–320

Abstract

Large-scale dieback of ash trees (*Fraxinus* spp.) caused by the fungus *Hymenoscyphus fraxineus* is posing an immense threat to forest health in Europe, requiring effective monitoring at large scales. In this study, a pipeline was created to find ash trees and classify dieback severity using high-resolution hyperspectral imagery of individual tree crowns (ITCs). Hyperspectral data were collected in four forest sites near Cambridge, UK, where 422 ITCs were manually delineated and labelled using field-measurements of species and dieback severity (for ash trees). Four algorithms, namely linear discriminant analysis (LDA), principal components analysis coupled with LDA (PCA-LDA), partial least squares discriminant analysis (PLS-DA) and random forest (RF), were used to build classification models for species and dieback severity classification. The effect of dark-pixel filtering on classification accuracy was evaluated. The best performing models were then coupled with automatic ITC segmentation to map species and ash dieback distribution over 16.8 hectares of woodland. We calculated and partitioned the coefficient of variation (CV) of the reflected ash spectra to find variable wavebands associated with dieback. PLS-DA and LDA were most accurate for classifying ITC species identifies (overall accuracy >90%), whereas RF was most accurate for classifying ash dieback severity (overall accuracy 77%). Dark pixel filtering further increased the accuracy of species classification (+6%), but not disease classification. The reflectances of narrow blue (415 nm), red-edge (680 nm) and NIR (760 nm) bands had high CV across disease classes and should be included if multispectral imagery were to be used to monitor ash dieback. The study demonstrates the possibility of using remote sensing to forward epidemiological research by monitoring forest pathogens in landscape scales, which would allow temperate forest managers to control pathogen outbreaks, assess associated impacts and restore affected forests much more effectively.

Introduction

Ash dieback currently poses an immense threat to forest health in continental Europe and the British Isles (Baral et al., 2014; McKinney et al., 2014). Caused by the ascomycete *Hymenoscyphus fraxineus* introduced from Far East Asia, ash dieback has quickly spread in Europe since its 1992 discovery in Poland (Drenkhan et al., 2017; McKinney et al., 2014). Over 95% of European ash trees, including the common European ash (*Fraxinus excelsior*), are susceptible to infection (Enderle et al., 2019; McKinney et al., 2014). The spread of *H. fraxineus* is mainly caused by airborne ascospores, which germinate and enter leaves through stomata, causing lesions to develop on leaves, rachis and branches, eventually leading to crown and tree death (Mansfield et al., 2018). The epidemic may have serious repercussions for ash-dependent biota in Europe, including at least 74 UK invertebrate species that are associated with *F. excelsior* (Littlewood et al., 2015). Efforts to mitigate the impacts of the pathogen focus on finding trees with resistant genetic markers in natural populations of ash such that resistant trees can be planted back in affected forests (Pautasso et al., 2013). Consequently, there is an urgent need to develop effective

methods to identify healthy and diseased ash trees at the landscape-scale, which will help epidemiologists and forest managers effectively model, control, assess and mitigate the epidemic.

In recent years, an increasing number of studies have demonstrated the effectiveness of passive remote sensing for monitoring forest pathogens over large areas, even identifying infection even before visible symptoms appear (Pietrzykowski et al., 2007; Waser et al., 2014; Zarco-Tejada et al., 2018). The accuracy and specificity of detection achieved in a particular classification task depends mainly on the resolution of images collected, type of imagery used and analysis methods (Stone & Mohammed, 2017). Low-resolution multispectral imagery, such as MODIS and Landsat, has been used for coarse-scale detection of disease outbreaks (Housman et al., 2018), but are unsuitable for detecting ash dieback. Ash trees usually grow in mixed-species forests, where hetero-specifics are also affected by a range of pests and pathogens, including Dutch elm disease and acute oak decline (Camilo-Alves et al., 2017; Karnosky, 1979), making it necessary to monitor ash dieback at the scale of individual tree crowns (ITCs). More accurate results are likely to be produced using high-resolution remote sensing data, as hetero-specifics can be filtered out, through species classification, before assessing ash dieback. Another important aspect of such data is the spectral resolution of the imagery. In particular, hyperspectral imagery, which measures the reflectances of hundreds of narrow wavebands, captures fine biochemical details of plant foliage (Asner et al., 2015), providing more information for accurate plant pathogen detection than multispectral or red-green-blue (RGB) imagery that measure reflectance in only a few broader wavebands (Stone & Mohammed, 2017; Zarco-Tejada et al., 2018). Lastly, data processing and analysis often greatly affect classification results (Calderón et al., 2015; Fassnacht et al., 2016; Waser et al., 2014; Zarco-Tejada et al., 2018). For example Asner et al. (2015) showed that discarding dark pixels from hyperspectral images before model-building improves the accuracies of classification tasks, and Lu and Weng (2007) reviewed studies on image classification and found the choice of model-building algorithms to have significant effects on model performance.

Currently, no study has used hyperspectral imagery to monitor *H. fraxineus* in mixed species forests in Europe. We know of only one study that used remote sensing data to develop models to monitor fungal ash dieback detection (Waser et al., 2014). The study used pan-sharpened satellite multispectral images to detect ash dieback and compared the performance of several parametric algorithms, including principal components analysis (PCA) and linear discriminant analysis (LDA) (Waser et al., 2014). Another study on emerald ash borer in North America suggests the benefits of using hyperspectral imagery in similar classification tasks (Pontius et al., 2008). Newer machine learning algorithms, such as random forest (RF), which have improved accuracies in many forest classification tasks in previous studies (Fassnacht et al., 2016; Stone & Mohammed, 2017), have never been applied to fungal ash dieback detection. Additionally, the effect of dark-pixel filtering on species and disease classification accuracy has not been evaluated. Lastly, a systematic assessment of the wavebands important for disease classification is also currently missing, despite previous work suggesting that some regions of the spectrum, such as the red-edge, might be more relevant to stress than others (Horler et al., 1983).

This study aims to fill these research gaps by developing a method to monitor ash dieback using high-resolution hyperspectral imagery and achieve four major goals. First, four different algorithms, including linear discriminant analysis (LDA), principal components analysis coupled with LDA (PCA-LDA), partial least squares differential analysis (PLS-DA) and random forest (RF), are tested for their ability to classify tree crowns according to species and ash crowns according to ash dieback severity. Second, the study evaluates whether dark-pixel filtering before model-building improves the accuracy of species and disease classification. Third, the study combines species and ash dieback classification models with automatic ITC segmentation to map species and ash dieback distribution over thousands of trees. This tests whether the models could produce species and dieback maps at a landscape scale that are useful for epidemiological research. Fourth, by partitioning the variation in the reflected spectra of ash trees, the study identifies specific wavebands important for less costly multispectral disease detection. By doing so, we provide a new approach for monitoring ash dieback in Europe that advances the application of remote sensing in plant epidemiology.

Materials and Methods

Study site

The study was carried out in four forests near Cambridge, United Kingdom, namely Madingley (UTM 31 N 298 443 mE 5 789 254 mN), Hayley (UTM 30 N 287 029 mE 5 783 351 mN), Gamlingay (UTM 30 N 282 142 mE 5 784 199 mN) and Bradfield woods (UTM 31 N 351 568 mE 5 783 559 mN). A map showing the locations of the four sites can be found in Figure 1. The vegetation of the four sites is mainly comprised of temperate broadleaf deciduous species. Dominant species in the canopy include oak (*Quercus robur*), ash (*Fraxinus*

excelsior), hazel (*Corylus avellana*) and field maple (*Acer campestre*). There are also notable presences of elm (*Ulmus* spp.) in Madingley, common alder (*Alnus glutinosa*) in Bradfield and silver birch (*Betula pendula*) in both Hayley and Bradfield. All four forests have a long history of management, creating a diverse mixture of old growth and secondary forests (Bellamy et al., 2009; Rackham, 1975; Rackham & Coombe, 1996). At the time of the survey, the forests were affected by ash dieback to different extents. Among the four sites, Bradfield wood was most severely affected; Madingley and Hayley woods consist of a mixture of diseased and healthy trees; while ash trees in Gamlingay were at early stages of infection. The wide range of dieback severity makes the four study sites ideal for studying ash dieback detection.

Field data

Field data were collected between July and September of 2018. In total, 422 trees were sampled, including 292 ash trees and 130 trees of other species (Table S1). Only overtopping trees clearly visible in high-resolution airborne RGB images were selected. Ash trees were chosen to represent the full spectrum of dieback severities. Following the methods described in Pontius et al. (2008) and Waser et al. (2014), the chosen ash trees were diseased scored from 0 (symptomless) to 10 (dead) according to the percentage of the crown showing symptoms. The scores were later collapsed into a more robust three-class system to increase the sample size, with trees scored 0–2 classified as healthy, 3–5 as infected and 6–10 as severely infected. Healthy trees (0–2) had lush foliage with occasional dead branches that could not be confidently attributed to fungal ash dieback; infected trees (3–5) had clear signs of infection, but still possessed a closed canopy; while heavily infected trees (6–10) developed open canopies as a result of fungal ash dieback. The location of each crown was recorded using a differential GPS (Geneq SXBlue II+) when reported geopositioning errors were <2 m.

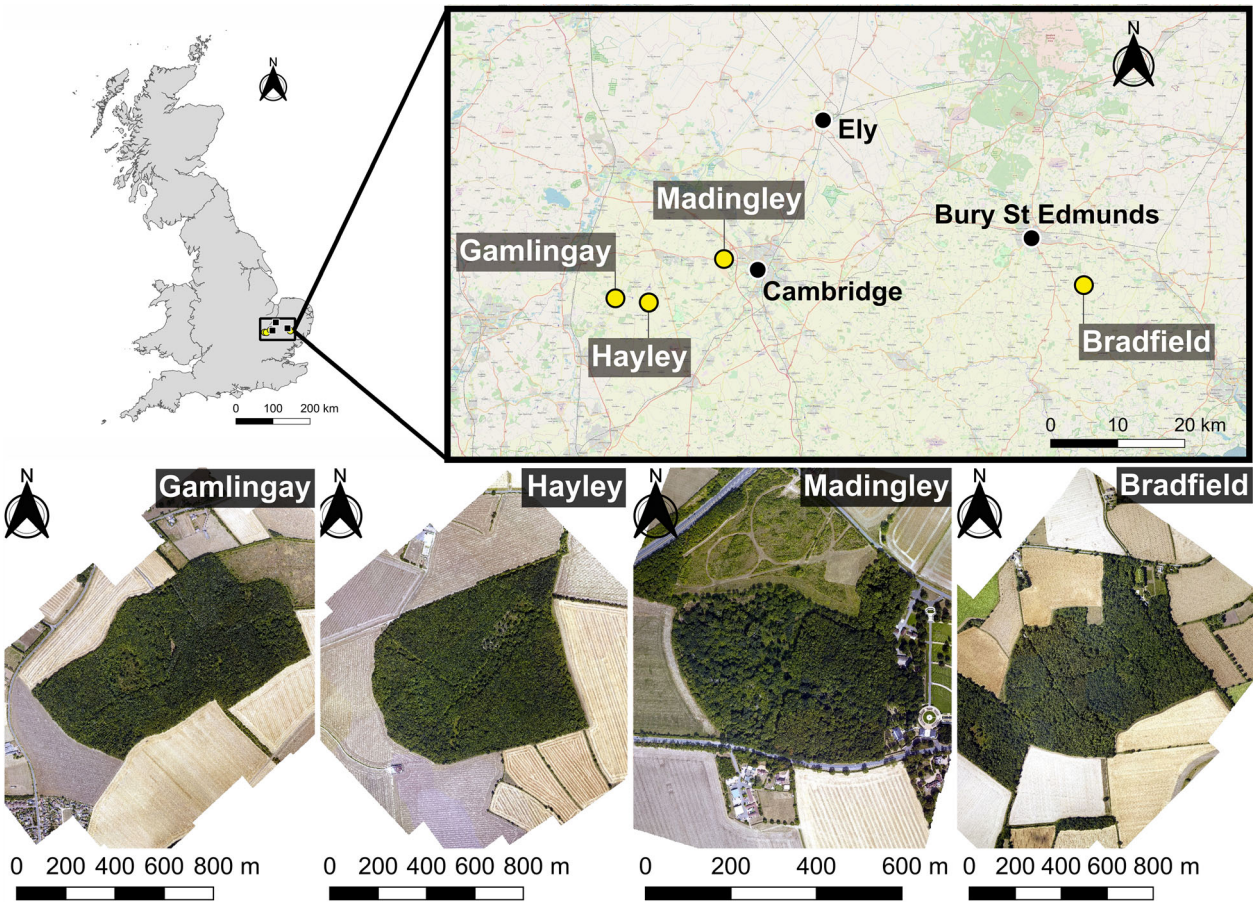

Figure 1. The RGB imagery and location of the four forest study sites. The top left panel shows the location of the sites relative to the outline of Great Britain (Office of National Statistics, 2019). The top right panel shows the location of the sites relative to the Cambridge (OpenStreetMap contributors, 2019).

Remote sensing data

Remote sensing data were collected by 2Excel geo by a manned aircraft (Piper PA-31 Navajo) flown across the four study sites over two clear days in August 2018, when deciduous trees had lush foliage. The aircraft was equipped with a Norsk Elektro Optikk (NEO) hyperspectral camera (Hyspex VNIR 1800) with a field of view of 17°. The hyperspectral data include reflectance at 186 bands spanning across the visible and near-infrared regions of the spectrum (410–1001 nm) with a spectral resolution of 3.26 nm and a spatial resolution of 0.32 m. QUick Atmospheric Correction (QUAC), designed for images collected over terrestrial systems, was used to remove atmospheric effects from the dataset (Bernstein, 2012). To prevent differences in illumination level across pixels from confounding the analysis, the hyperspectral data were normalized by dividing the reflectances with the average reflectance of all wavebands in the same pixel, as suggested by Wu (2004). Additionally, 66 vegetation indices (VIs), including both broadband and narrowband indices, were calculated from the hyperspectral data based on formulae in the Index DataBase (IDB) (Henrich et al., 2009) (Table S2). The 66 VIs together with the reflectances of the 186 wavebands yielded 252 predictor variables, which were used to train the classification models. The aircraft also carried a Phase One Industries iXA80 camera, which collected RGB images of the sites with a spatial resolution of 0.06 m. Using SimActive Correlator3D, a photogrammetry software, digital surface models (DSMs) of the four sites with a ground resolution of 0.34 m were generated from the RGB images. For Madingley, the digital terrain model (DTM) was available from the UK government (https://data.gov.uk/dataset/6a117171-5c59-4c7d-8e8b-8e7aefe8ee2e/lidar-composite-dtm-2017-1m). By subtracting the DTM from the DSM, a canopy height model (CHM) representing tree heights was created for the site.

Remote sensing data corresponding to the species-identified and disease-scored trees listed in Table S1 were extracted from manual delineations of the crowns. The field-measured GPS points of the 422 surveyed trees (Table S1) were projected on the RGB imagery using QGIS. The laptop, which displayed the data, was then brought to the field. With the GPS points serving as a rough guide, and by discerning the shapes and colours of nearby crowns, crowns on the displayed RGB imagery were matched with surveyed trees (Table S1). These crowns were manually delineated on site with the 'Add polygon feature' function in QGIS (Fig. 2). The hyperspectral data in pixels encircled by the polygons were extracted for further analysis.

Building models for species and ash dieback detection

For species classification, four algorithms were tested, including linear discriminant analysis (LDA), principal components analysis coupled with LDA (PCA-LDA), partial least squares discriminant analysis (PLS-DA) and random forest (RF). As relatively few birch and alder trees occurred within the study sites, the analysis focused on the five species with >20 crowns (ash, oak, hazel, field maple and elm). Nevertheless, interested readers can refer to Figure S1 for the effect of including birch and alder, as well as varying sample sizes, on the accuracy of one of the species classification models (PLS-DA). For the five main species, hyperspectral data from the manually delineated crowns were split into training (70%) and validation (30%) datasets. Care was taken to ensure that validation was done across tree crowns rather than across pixels that may share the same crown. LDA, PCA-LDA and PLS-DA models were trained at the pixel level, where models returned species identity predictions of individual pixels. For these models, tree crown level predictions were made by the majority voting of classified pixels. Due to computer (16 GB RAM) memory limitations, the RF model was created at the tree crown level, trained and validated by hyperspectral data averaged across the tree crowns. All model training was carried out in R (R Core Team, 2020).

LDA, PLS-DA and RF were also used to build models to classify ash crowns according to dieback severity based on hyperspectral data. As symptoms of ash dieback often manifest at the sub-crown level, a diseased tree may have a mixture of healthy and diseased branches. The resolution of the hyperspectral imagery was not high enough to reveal individual diseased branches, so it would be inappropriate to build disease classification models at the pixel level. Hence, the three-class (healthy, infected and severely infected) ash dieback models were trained and validated using hyperspectral data averaged across tree crowns. Among all ash crowns, 16 trees from each disease class were selected to comprise the validation dataset. The validation was mainly done across sites: training tree crowns were selected from Madingley and Hayley, while validation tree crowns were mainly selected from Bradfield and Gamlingay (with the exception of several trees in the 'infected' class, as the two sites have <16 delineated trees from this class).

Four measures of accuracy were used to evaluate model performance, namely overall accuracy, average user's accuracy, average producer's accuracy and the Kappa statistic. Detailed definitions are described by Fung and LeDrew (1988).

Dark pixel filtering

To test whether dark pixel filtering improved the accuracies of the species and ash dieback classification models created in this study, the average reflectance was calculated for each pixel in the hyperspectral dataset. The best species and ash dieback classification models were trained repeatedly with 0% to 90% of the darker pixels filtered out from the training dataset. The effects of such filtering on the performance of the models (overall accuracy) were recorded.

Species and ash dieback mapping

The species and ash dieback classification models with the highest accuracies were coupled with automatic individual tree crown segmentation to map the structure of the 16.8 ha Madingley forest (Rackham & Coombe, 1996), a site with a readily available CHM for segmentation. Automatic segmentation was performed using an adaption of itcSegment (Dalponte, 2016) (https://github.com/swinersha/Tree-crown-segmentation), which runs faster and is less memory-demanding compared to the original code, creating a shapefile layer containing polygons that encircled individual tree crowns. Polygons were classified to species by applying the most accurate species classification model to the encircled hyperspectral pixels, followed by the majority voting of classified pixels. Ash polygons were then classified by dieback severity using the disease model that produced that highest classification accuracy from Section 2.4. This produced two maps showing tree species distribution and ash dieback severity over the entire site respectively.

Decomposing hyperspectral data for multispectral ash dieback detection

Hyperspectral data are costly to collect and process (Stone & Mohammed, 2017). Thus, identifying the wavebands important for multispectral ash dieback detection could substantially reduce the cost of ash dieback monitoring. To identify these wavebands, the reflected spectra of ash crowns in the three disease classes were plotted out. Additionally, the coefficients of variation (CV) were calculated for each waveband for (1) pixels within individual tree crowns, (2) ash crowns within the same disease class and (3) groups of ash crowns of different disease classes.

Results

Species and ash dieback classification accuracy

Two supervised parametric algorithms, namely partial least squares discriminant analysis (PLS-DA) and linear discriminant analysis (LDA), produced the most accurate species classification models; while random forest (RF), a non-parametric machine learning algorithm, most accurately classified ash dieback severity (Table 1).

For species classification, before filtering out dark pixels, models created by both LDA and PLS-DA achieved ≥90% in three measures of tree-crown-level accuracy, with an associated Kappa of 0.877, which indicates a near-perfect agreement between predicted and actual species classes after accounting for the effect of random chance (Landis & Koch, 1977). The confusion matrix of the best-performing PLS-DA model (Table 2) demonstrates the ability of the model

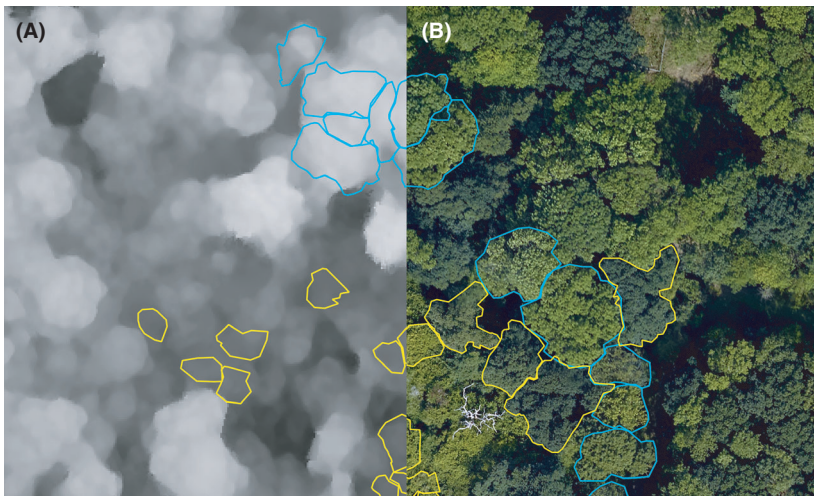

Figure 2. Manual tree crown delineations in Madingley woods with (A) the canopy height model (CHM) and (B) the high resolution RGB image in the background. Blue delineations bound ash tree crowns, while yellow delineations bound hetero-specifics.

Table 1. Accuracies achieved by species and ash dieback severity classification models before dark-pixel filtering.

Task	Method	Overall accuracy	Average user's accuracy	Average producer's accuracy	Kappa
Species	LDA	**0.902** [**0.853**]	0.932 [0.857]	0.900 [**0.846**]	**0.877** [**0.803**]
	PCA-LDA	0.878 [0.741]	0.933 [0.712]	0.875 [0.721]	0.846 [0.653]
	PLS-DA	**0.902** [0.846]	**0.954** [**0.861**]	**0.921** [0.804]	**0.877** [0.790]
	RF	0.756	0.764	0.751	0.693
Ash dieback severity	PLS-DA	0.646	0.680	0.646	0.469
	LDA	0.562	0.546	0.562	0.344
	RF	**0.771**	**0.772**	**0.771**	**0.656**

Algorithms used to build the models include linear discriminant analysis (LDA), principal component analysis coupled with LDA (PCA-LDA), partial least squares discriminant analysis (PLS-DA) and random forest (RF). The unbracketed number represents classification accuracies on a tree crown level, whereas the numbers in square brackets represent classification accuracies on a pixel level. The numbers in bold indicates the highest accuracies achieved.

Table 2. Confusion matrix obtained from the PLS-DA species classification model showing predicted and actual species IDs.

		Prediction					
		Ash	Elm	Field maple	Hazel	Oak	**Producer's accuracy**
Reference	Ash	**10 [9379]**	0 [58]	0 [21]	0 [31]	0 [105]	1 [0.978]
	Elm	1 [422]	**6 [2797]**	0 [53]	0 [77]	0 [345]	0.857 [0.757]
	Field maple	2 [1798]	0 [40]	**6 [2926]**	0 [79]	0 [197]	0.75 [0.581]
	Hazel	0 [287]	0 [19]	0 [116]	**7 [2052]**	0 [40]	1 [0.816]
	Oak	0 [440]	0 [302]	0 [416]	0 [13]	**8 [9484]**	1 [0.890]
User's accuracy		0.769 [0.761]	1 [0.870]	1 [0.828]	1 [0.911]	1 [0.932]	

The unbracketed figures represent number of tree crowns and associated classification accuracies, while figures in square brackets represent the number of pixels and corresponding classification accuracies. The figures corresponding to correctly classified crowns are bolded. In summary, the model achieved an overall accuracy of 0.902 [0.846], an average user's accuracy of 0.954 [0.861], an average producer's accuracy of 0.921 [0.804], and a Kappa statistic of 0.877 [0.790] before dark pixels were filtered.

to accurately predict crown species identity for all five tested species, achieving >75% crown classification accuracies even for species pairs with similar spectral signatures, such ash and field maple (Fig. S2). Performing principal components analysis (PCA) as an additional dimensionality reduction step before LDA failed to improve species classification accuracy. The poor performance of PCA was highlighted by the poor species clustering on the plots between principal components (PCs) when compared with equivalent latent variable (LV) plots generated by PLA-DA (Figs. S3 and S4). Random forest (RF) was also unable to generate a model with comparable classification accuracies, possibly since the model was not trained on a pixel level due to computer memory limitations.

In disease classification, random forest (RF) produced the model that most accurately classified ash crowns into three severity classes, achieving ≥75% in the three measures of accuracy. The model also had a Kappa of 0.656, which indicated substantial agreement between model predictions and the reference after accounting for chance (Landis & Koch, 1977). The confusion matrix generated while validating the RF model (Table 3) demonstrates that severely infected trees were well-identified in the model, with all 16 severely infected trees correctly classified. Less severely infected trees were more difficult to identify, but still correctly classified in most cases, with the model achieving 62-69% producer's accuracies and >75% user's accuracies for these ash crowns.

Table 3. Confusion matrix generated by classifying ash trees into three disease classes using a random forest (RF) ash dieback classifier.

		Prediction			
		Healthy	Infected	Severely infected	**Producer's accuracy**
Reference	Healthy	**11**	3	2	0.688
	Infected	3	**10**	3	0.625
	Severely infected	0	0	**16**	1
User's accuracy		0.786	0.769	0.762	

The model achieved an overall accuracy of 0.771, an average user's accuracy of 0.772, an average producer's accuracy of 0.771 and a Kappa statistic of 0.656. The numbers correspond to the number of tree crowns in the validation dataset falling into each category. The numbers representing the correctly classified crowns are bolded.

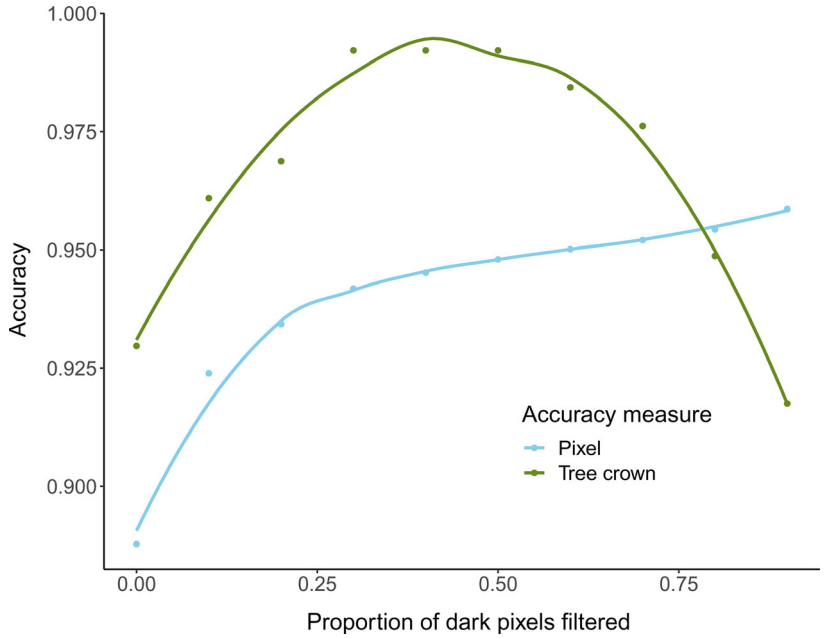

Figure 3. Change in overall species classification accuracies of PLS-DA models on a tree crown and pixel levels in response to dark pixel filtering.

Improved species classification but poorer disease detection through dark pixel filtering

Dark pixel filtering substantially improved the overall accuracy of the PLS-DA species classification model (Fig. 3). On the pixel level, the classification accuracy continuously increased as larger proportions of dark pixels were filtered out. The highest pixel classification accuracy was achieved when 90% of the darker pixels were filtered, with no sign of peaking. In contrast, dark pixel filtering increased tree crown level species classification accuracy initially until 50% of the darker pixels were filtered out, after which it decreased. Filtering approximately 40% of the darker pixels optimized the species classification accuracy, increased by 6.3% when compared to no filtering (Fig. 3).

Contrary to species classification, dark pixel filtering did not improve ash dieback severity classification of the best-performing RF model. Rather, dark pixel filtering led to a general decline in overall accuracy in disease classification, especially for healthy and heavily infected ash tree crowns (Fig. 4), suggesting a significant role of dark pixels associated with gaps in tree crowns in indicating dieback severity.

Mapping ash dieback in forests

Individual tree crown (ITC) segmentation algorithms, coupled with the best-performing PLS-DA species classification model and RF disease classification model, produced detailed maps of ash and ash dieback distributions. Figure 5A shows the general structure of the site (Madingley wood) based on Rackham and Coombe (1996), who provided a detailed account of the species composition of the wood, along with an inset demonstrating the slightly over-segmented crown polygons generated by the modified itcSegment package (Dalponte, 2016) from the CHM. Visual inspection revealed that, of the 383 polygons overlapping with manually delineated crowns, 67% did not significantly (>30%) encircle neighbouring trees, which is acceptable considering that many densely packed trees, especially coppiced hazel, were difficult to disentangle even in the field. Species mapping by the PLS-DA model revealed the secondary forest to the east being dominated by ash, old-growth forest to the west composed mainly of oak, the southwest corner housing a large colony of elm and hazel trees appearing in coppice plots that were not overtopped by other species (Fig. 5B). The mapped species distribution is consistent with that described by Rackham and Coombe (1996). Applying the RF disease classification model on tree crown classified as ash in Figure 5B revealed the distribution of ash dieback in the site (Fig. 5C). In general, ash trees that were on the edge of the forest were found to be more heavily infected than other trees in the site.

Evaluating the need of using hyperspectral data in disease classification: The reflected spectra and coefficient of variation (CV) of ash trees

To evaluate which hyperspectral wavebands were necessary for accurate disease classification, the reflected spectra of ash crowns with different disease severities

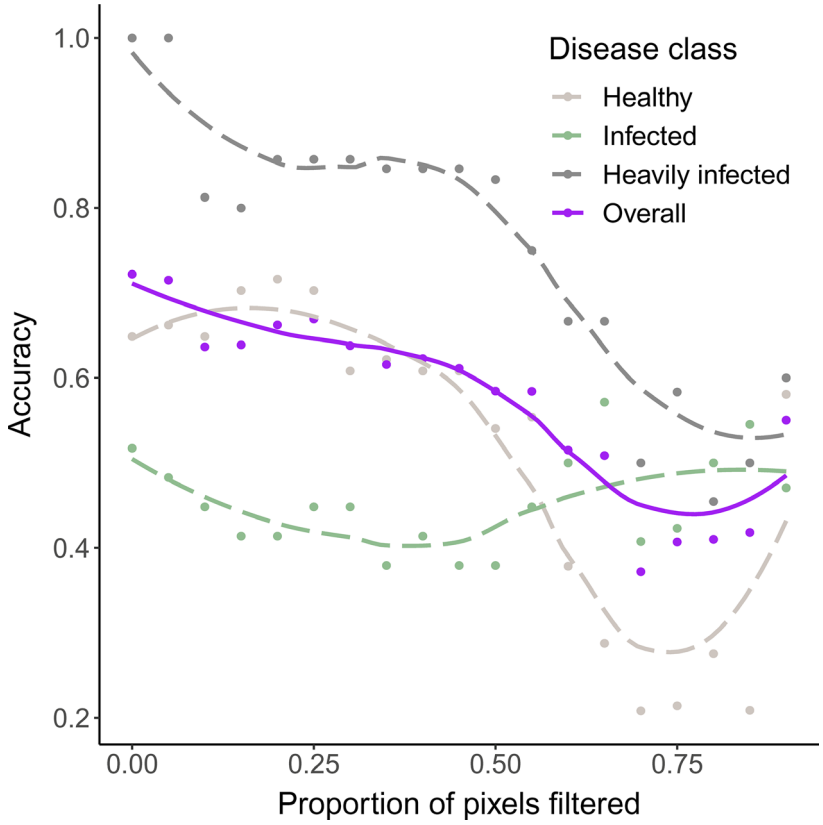

Figure 4. The changes in accuracy of classification of the random forest (RF) disease severity classifier over an increase proportion of dark pixels filtered from the training and validation datasets. The purple solid line corresponds to overall accuracy. The dotted lines correspond to the producer's prediction accuracies of trees in the three disease classes.

(Fig. 6A) and partitioned coefficients of variation (CV) of the hyperspectral data (Fig. 6B) were plotted. More heavily infected ash crowns were found to have higher red-edge (around 680 nm) and lower NIR (750 nm–1000 nm) reflectances (Fig. 6A). With most of the variation in reflectance occurring among pixels within the same crown, CV between crowns with different disease classes only accounted for a small fraction of the total variation. A small number of wavebands near the blue (around 415 nm), red-edge (around 680 nm) and near-infrared (around 760 nm) regions were found to be particularly variable across disease classes. These sharp CV peaks (Fig. 6B) in these regions show the importance of including narrow wavebands and narrow-band vegetation indices (VIs), rather than broad-band features that obscures detailed spectral signatures, as predictors when building disease classification models.

Discussion

Performance of different model training algorithms

As suggested by previous literature, the approach used to build the species classification model greatly affected classification accuracy (Fassnacht et al., 2016; Waser et al., 2014). The four approaches used in this study have been widely used in the past for building models based on data with high dimensionality, each with its own reported advantages and shortfalls (summarized in Table 4). Here, we found supervised parametric approaches (LDA and PLS-DA) outperforming PCA-LDA. This contradicts previous studies that have repeatedly suggested that adding an extra PCA dimensionality reduction step before LDA improves LDA performance (Downey, 1994; Kemsley, 1996; Kher et al., 2006). Our results, however, agree with Waser et al. (2014), who also found PCA-based approaches to perform worse than LDA in ash dieback detection using multispectral satellite images. Waser et al. (2014) did not provide a thorough explanation for why results contradicted with other studies, but based on the poor species clustering in the graph showing PC2 against PC1 (Fig. S4), it is possible that the unsupervised nature of PCA failed to select features that were relevant to species classification when most of the variance lay within tree crowns and were unrelated to species differences (Fig. S3). The accuracy of RF in species classification was also lower than the two supervised parametric approaches (LDA and PLS-DA). That was likely because the RF model was trained and validated at a tree

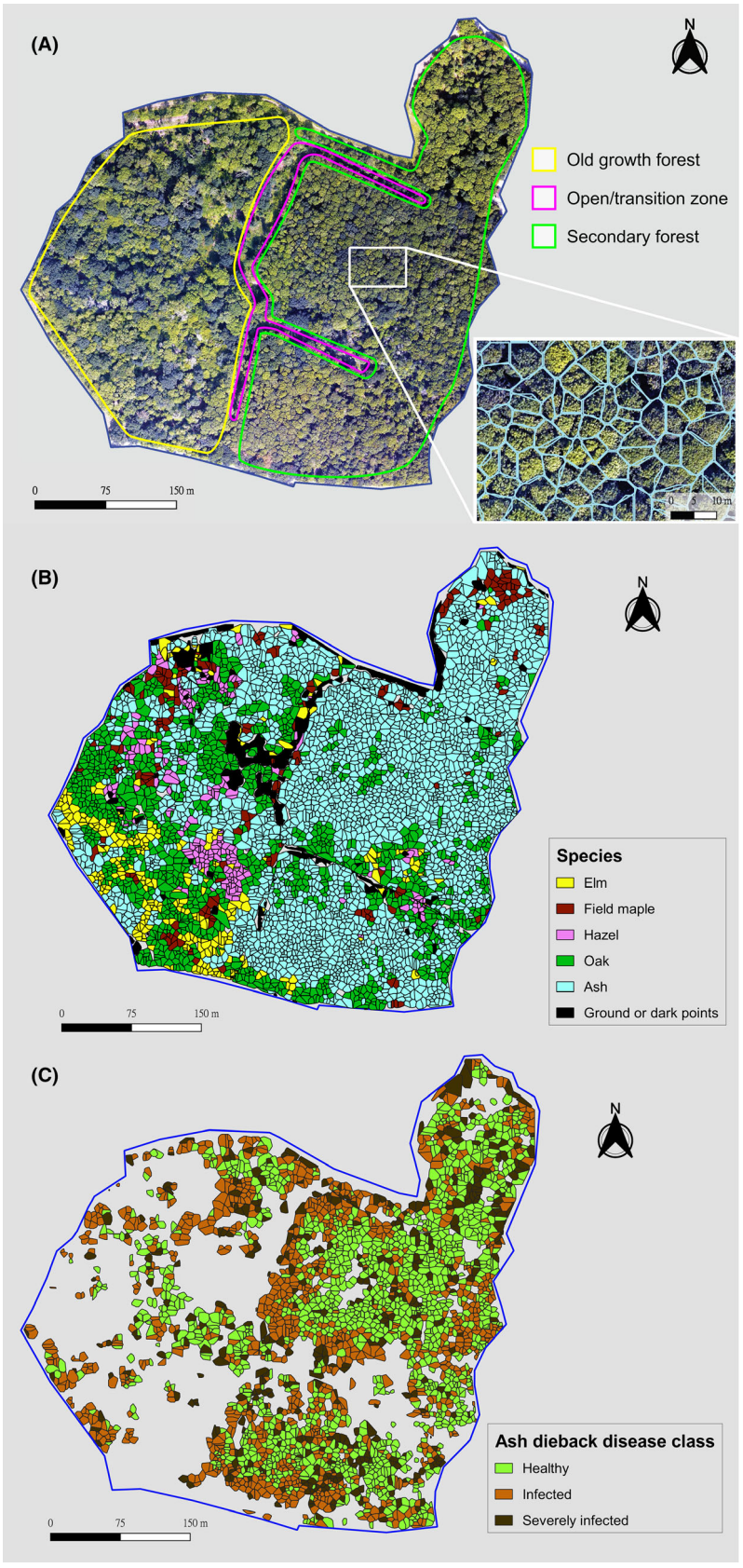

Figure 5. Mapping the distribution of tree species and ash dieback severity in the 16.8 ha Madingley wood. (A) shows the general structure of the site as described by Rackham and Coombe (1996), with the inset demonstrating polygons generated by a modified version of itcSegment. (B) shows the PLS-DA species map. (C) shows the RF dieback disease classmap.

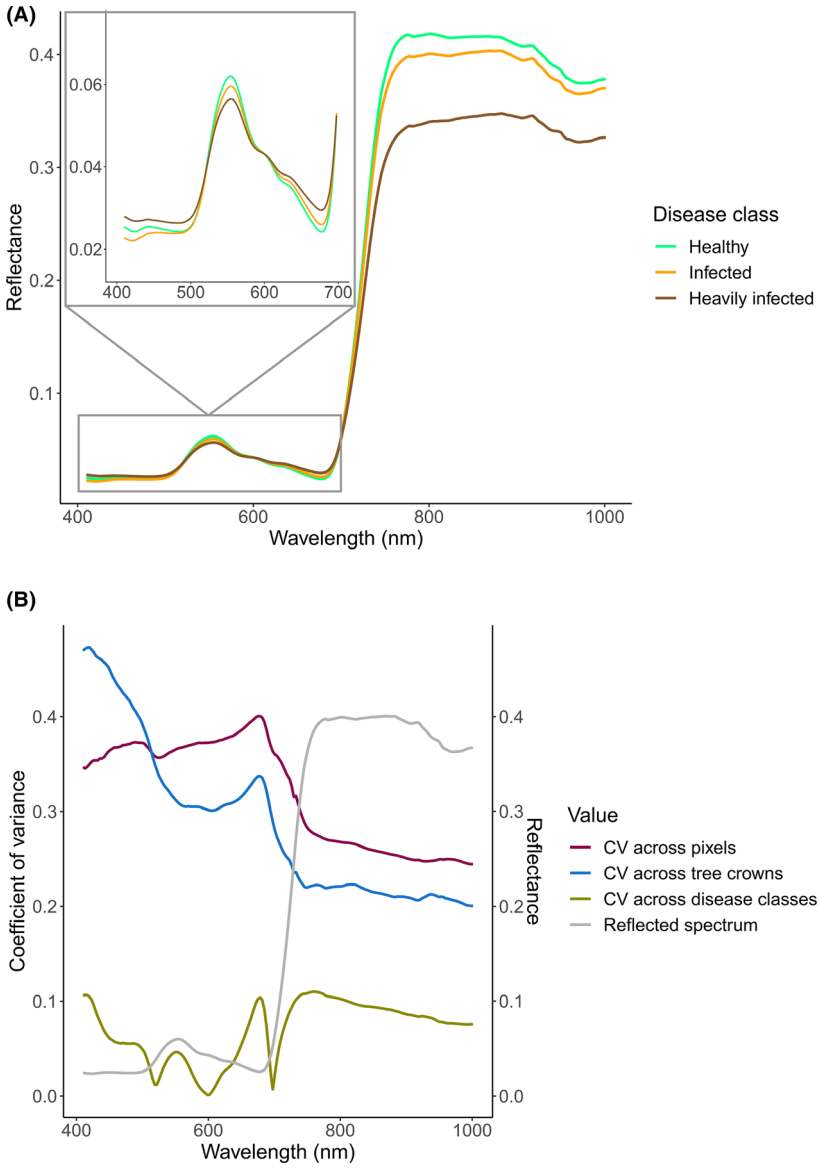

Figure 6. Graphs showing (A) the mean reflected spectra of ash trees with different disease severities and (B) the associated coefficient of variation (CV) of the reflectances across the observed spectrum from visible to NIR [409–1001 nm]. In (A), the inset shows reflectance in the visible [409–700 nm] region. In (B), the CV was partitioned into the variation among pixels within the same tree crown (purple), the variation among crowns carrying the same disease score (blue), and the variation among groups of crowns with different disease scores (yellow).

crown rather than pixel level due to computer memory limitations. While this highlights specific limitations in using machine learning algorithms in large scale classification tasks, the accuracies achieved were not directly comparable. In disease classification, where all three models (RF, LDA and PLS-DA) were trained and validated on tree-crown-averaged hyperspectral data, RF clearly outperformed the two supervised parametric algorithms. This demonstrates that machine learning algorithms are best at generalizing complex, non-parametric relationships between the ash dieback severity and hyperspectral reflectances, and possibly that between species and hyperspectral data if memory limitations were to be lifted. The accuracies achieved by RF may further benefit from steps to address multicollinearity. Previous work has found RF models to be less stable and less transferrable when predictors were highly intercorrelated (Toloşi & Lengauer, 2011; Weaving et al., 2019). This may have affected the accuracies of the RF disease classification model validated across sites and flightlines. An evaluation of different preemptive steps to reduce the effect of multicollinearity on RF, such as dimensionality reduction and group selection algorithms, may be worth exploring in the future.

The >90% overall accuracy and >0.85 Kappa statistic achieved by PLS-DA and LDA species classifiers produced in this study are higher than that reported by most similar studies, especially after dark pixel filtering, which further boosts the overall PLS-DA classification accuracy to >95%.

Table 4. A comparison on the advantages and disadvantages of the four different approaches used to build species classification models.

Method	Advantages	Disadvantages
Linear discriminant analysis, a supervised, parametric approach LDA	Simple and quick to build Have been successfully implemented on remote sensing data for species classification and ash dieback detection in the past (Waser et al., 2014)	Classification accuracies and the stability of the model affected by collinearity of predictors (Naes & Mevik, 2001) Less powerful than non-parametric approaches in using more complex patterns in classification
Principal component analysis followed by linear discriminant analysis, an unsupervised (PCA), parametric approach	Simple and quick to build; PCA is a widely used multivariate statistical technique (Abdi & Williams, 2010) Past studies have suggested increased classification accuracies with PCA as an extra dimensionality reduction step (Downey, 1994; Kemsley, 1996; Kher et al., 2006) Unsupervised dimensionality is often considered to be less susceptible to overfitting (Abdi & Williams, 2010)	PCA is unsupervised PCA has been suggested to perform worse than LDA in building species and ash dieback classifiers using remote sensing data in previous work (Waser et al., 2014) Less powerful than non-parametric approaches in using more complex patterns in classification
Partial least squares discriminant analysis, a supervised parametric approach	Makes no assumption of the distribution of the data, hence more flexible than other parametric discriminant algorithms (Lee et al., 2018) Less memory intensive than RF and other deep learning approaches Robust against high collinearity (Lee et al., 2018; Naes & Mevik, 2001)	Slower than other parametric approaches due to its relative complexity Less powerful than non-parametric approaches in using more complex patterns in classification
Random Forest; a supervised, non-parametric, ensemble machine learning approach	The non-parametric nature of RF makes it suitable for finding complex patterns in datasets Faster than many other deep learning approaches Previous studies in the remote sensing community suggest that RF, as an ensemble approach, achieves higher accuracies than other non-parametric supervised classifiers such as Classification and Regression Tree (CART) or Support Vector Machine (SVM) (Belgiu & Drăguţ, 2016)	Compared with the other three parametric approaches, RF is slow and memory intensive when applied over large datasets Non-parametric approaches are often considered to be more susceptible to overfitting, reducing applicability outside the training window (Belgiu & Drăguţ, 2016)

For comparison, Hill et al. (2010) attempted to classify six tree species in Monks Wood, Cambridgeshire using a single multispectral image. Using 165 trees for training, the study reported a classification accuracy of 71% and a Kappa of 0.63. Dalponte et al. (2012) reported overall and Kappa accuracies of 83% and 0.77, respectively, in an attempt to classify trees in the Italian Alps into 8 classes (7 species classes +non-forest) based on reflectances of selected waveband and LiDAR data. Fassnacht et al. (2016) reviewed 101 studies attempting species classification and found overall accuracies of 80-85% depending on the type of data used. The high accuracies achieved here to demonstrate the merit of using hyperspectral imagery with high spatial and spectral resolution in species classification tasks. It also shows that a small number of accurately manually delineated tree crowns might be preferable to a large number of poorly delineated crowns in training species classification models. Lastly, it highlights the importance of preprocessing the hyperspectral data, such as filtering away dark pixels before model-building, in improving classification accuracies.

The ash dieback classification accuracies achieved by the RF model in this study (77% overall accuracy and a 0.66 Kappa) also compares well with previous studies that attempted to detect pest infestations pathogen infections in forests, which commonly reported accuracies ranging from 70 – 90% (Hall et al., 2016; Murfitt et al., 2016; Waser et al., 2014). Given that Waser et al. (2014) also reported a 77% overall accuracy in classifying ash trees affected by ash dieback into four disease classes using satellite multispectral imagery, at a glimpse, the classification accuracies achieved in this study may seem underwhelming. However, it should be noted that Waser et al. (2014) used heavily infected German forests in their study, which contain many more large, severely infected trees than Madingley and Hayley woods, used to train the RF model in this study. Additionally, the RF model we built was validated mainly with trees from a different site, whereas Waser et al. (2014) validated the classification models by resampling the training dataset, which tends to produce slightly higher accuracies (Yadav & Shukla, 2016). In fact, when LDA, the preferred approach used by

Waser et al. (2014) to classify diseased ash trees, was applied on the dataset in this study, the overall accuracy was low (56% overall accuracy and a Kappa of 0.34) (Table 1), suggesting that trees included in this study were more challenging to classify. Based on the two studies alone, no clear conclusion could be made as to whether hyperspectral data outperforms multispectral data in ash dieback detection.

The effect of dark pixel filtering on species and disease classification

Dark pixel filtering before training and validating the PLS-DA species classification model significantly improved model accuracy. Yet, while species classification accuracy at the pixel level increased monotonically as more pixels were filtered, a parabolic curve with an optimum (40% pixels filtered) was observed for majority-voted classifications of tree crowns. These patterns demonstrated that brighter hyperspectral pixels contained more information for classifying species, echoing the results of Asner et al. (2015), who found that brighter pixels better predicted canopy traits. However, when a large proportion of pixels were filtered out, the majority voting on the remaining pixels no longer produced accurate classifications at the tree crown level, resulting in the subsequent drop in accuracy. This creates an interesting trade-off between data fidelity and quantity (Dehghani et al., 2017). The optimum proportion of pixels likely varies according to the number of pixels per object to be classified, which is a function of spatial resolution object size. We suggest that future studies should routinely carry out tests to find the optimum if dark pixel filtering is to be performed for similar classification tasks.

Intriguingly, dark pixel filtering was counterproductive in ash dieback classification and lowered the accuracy of the RF disease classification model. This showed that the widening of gaps within crowns associated with disease infection might have acted as an important proxy for the RF model to predict ash dieback severity. We also found the drop in classification accuracy associated with dark pixel filtering to be more pronounced for 'healthy' and 'severely infected' crowns, but less so for 'infected' crowns, suggesting that mildly infected trees still had relatively closed canopies and but exhibited spectral features associated with subtle changes in leaf biochemistry rather than an increased proportion of dark pixels associated with crown dieback. It is important to note that the reflected spectra of these dark, gap pixels in heavily infected ash crowns might be affected by the background vegetation, as one previous study on larch disease suggested (Barnes, 2018). If gaps were used by the RF model as a proxy for heavily diseased ash, the model may have to be adjusted before being applied to ash trees growing over other backgrounds.

Mapping species and ash dieback distribution

Applying the PLS-DA species classification and RF fungal ash dieback classification model onto one of the field sites (Madingley wood) with a readily available CHM for ITC segmentation produced maps of species and ash dieback distribution that agreed well with past literature (Rackham & Coombe, 1996). The maps revealed that ash trees on the forest edge exhibit more severe dieback symptoms. This may be due to confounding background vegetation (Section 4.2.), but it could also be caused by a range of biological factors. Edge effects are known to bring additional stress to temperate trees (Reinmann & Hutyra, 2017) or might be younger and more susceptible to dieback (Enderle et al., 2019; McKinney et al., 2014). Alternatively, trees at the edge might receive a larger load of airborne *H. fraxineus* spores (Mansfield et al., 2018). Regardless, the results demonstrate the potential applications of ash dieback mapping in understanding the epidemiology of the pathogen.

The potential of using multispectral data to monitor ash dieback

Narrow wavebands near the blue (around 415 nm), red-edge (around 680 nm) and near infra-red (NIR) (around 760 nm) were found to have a high coefficient of variation (CV) across disease classes. As narrow bands near the red-edge are long known to be affected by chlorophyll status and plant stress (Horler et al., 1983), it is unsurprising that wavebands near the red-edge also vary according to disease severity (Fig. 6B). The reflectance of wavebands in the blue region are traditionally less known to be leaf stress indicators (Carter, 1993), but still exhibit large CVs among groups of different disease scores (Fig. 6B), possibly due to the manifestation of dead branches and exposure of the background vegetation associated with dieback. The significantly lower NIR reflectances among diseased ash trees have also been reported for hedgerow ash trees in the UK (Barnes, 2019). As NIR light is not absorbed by healthy plants but reflected to prevent overheating (Knipling, 1970), the pattern can be explained by ash trees suffering from dieback having less foliage, fewer branches and more open crowns. While results from this study do not preclude the possibility of achieving similar disease classification accuracies using multispectral data, which can be collected at lower costs, our results suggest the inclusion of narrow bands near 415 nm, 680 nm and 760 nm if possible.

Conclusion

The study showed that accurate species and ash dieback mapping can be attained by analysing hyperspectral data at the individual tree crown level. It also demonstrated how (1) collecting a dataset capturing the spectral features of interest, (2) choosing an effective model-building algorithm and (3) appropriately filtering dark pixels can further bolster classification accuracy. The pipeline described here can be scaled up to monitor forest pests and pathogens at a regional scale in Europe, not only for ash dieback, but also for other widespread forest pathogens such as Dutch elm disease (*Ophiostoma* spp.), chronic oak decline, acute oak decline and emerald ash borers (*Agrilus planipennis*). A large-scale object-based species-specific disease monitoring scheme based on remote sensing would be a key addition to forest epidemiology, allowing the detection of emerging pathogens, accurate transmission modelling and the screening of resistant genotypes. Forest managers would hence be able to swiftly stop emerging pathogens by deploying appropriate control measures, carry out impact assessments to estimate losses for specific outbreaks and mitigate the impacts by rehabilitating affected forests with resistant trees.

Acknowledgements

We acknowledge 2Excel geo for generously collecting and providing the high-resolution remote sensing data. We are grateful to Ross Hill for providing information on ash trees at Bradfield wood that his group had scored for disease in 2018, but did not end up being used for this study, and for helpful conversations when we were planning our work. We would also like to thank the University of Cambridgge, Suffolk Wildlife Trust, Wildlife Trust for Bedfordshire, Cambridgeshire & Nottinghamshire for providing support and access to the field sites. Additionally, the study would not have been possible without the valuable input from members of the Forest Ecology and Conservation Group, especially Dr Florian Zellweger, Dr Yi Zhang, Kyaw Sein Win Tun (O'Neill) and Clément Valle for their help with fieldwork; and Philip Sellars and Jonathan Williams for their advice on data analyses. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abdi, H. & Williams, L.J. (2010) Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**, 433–459. https://doi.org/10.1002/wics.101

Asner, G.P., Martin, R.E., Anderson, C.B. & Knapp, D.E. (2015) Quantifying forest canopy traits: imaging spectroscopy versus field survey. *Remote Sensing of Environment*, **158**, 15–27. https://doi.org/10.1016/J.RSE.2014.11.011

Baral, H.-O., Queloz, V. & Hosoya, T. (2014) *Hymenoscyphus fraxineus*, the correct scientific name for the fungus causing ash dieback in Europe. *IMA Fungus*, **5**, 79–80. https://doi.org/10.5598/imafungus.2014.05.01.09

Barnes, C. (2018) Remote sensing of larch disease and acute oak decline outbreaks in Britain. *University of Leicester*. https://hdl.handle.net/2381/42534

Barnes, C. (2019) Remote sensing technologies for the management of tree pests and diseases. *Quarterly Journal*, **113**, 115–120.

Belgiu, M. & Drăguţ, L. (2016) Random forest in remote sensing: a review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, **114**, 24–31. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.011

Bellamy, P.E., Hill, R.A., Rothery, P., Hinsley, S.A., Fuller, R.J. & Broughton, R.K. (2009) Willow Warbler *Phylloscopus trochilus* habitat in woods with different structure and management in southern England. *Bird Study*, **56**, 338–348. https://doi.org/10.1080/00063650902806914

Bernstein, L.S. (2012) Quick atmospheric correction code: algorithm description and recent upgrades. *Optical Engineering*, **51**, 111719. https://doi.org/10.1117/1.oe.51.11.111719

Calderón, R., Navas-Cortés, J. & Zarco-Tejada, P. (2015) Early detection and quantification of verticillium wilt in olive using hyperspectral and thermal imagery over large areas. *Remote Sensing*, **7**, 5584–5610. https://doi.org/10.3390/rs70505584

Camilo-Alves, C.S.P., Vaz, M., Da Clara, M.I.E. & Ribeiro, N.M.D.A. (2017) Chronic cork oak decline and water status: new insights. *New Forests*, **48**, 753–772. https://doi.org/10.1007/s11056-017-9595-3

Carter, G.A. (1993) Responses of leaf spectral reflectance to plant stress. *American Journal of Botany*, **80**, 239–243. https://doi.org/10.1002/j.1537-2197.1993.tb13796.x

Dalponte, M. (2016) *R Package'itcSegment': user manual.*

Dalponte, M., Bruzzone, L. & Gianelle, D. (2012) Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment*, **123**, 258–270. https://doi.org/10.1016/J.RSE.2012.03.013

Dehghani, M., Mehrjou, A., Gouws, S., Kamps, J. & Schölkopf, B. (2017). *Fidelity-weighted learning*. 6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc.

Downey, G. (1994) Tutorial review. Qualitative analysis in the near-infrared region. *Analyst*, **119**, 2367. https://doi.org/10.1039/an9941902367

Drenkhan, R., Solheim, H., Bogacheva, A., Riit, T., Adamson, K., Drenkhan, T. et al. (2017) *Hymenoscyphus fraxineus* is a leaf pathogen of local *Fraxinus* species in the Russian Far East. *Plant Pathology*, **66**, 490–500. https://doi.org/10.1111/ppa.12588

Enderle, R., Stenlid, J. & Vasaitis, R. (2019) An overview of ash (*Fraxinus* spp.) and the ash dieback disease in Europe. *CAB International*. https://doi.org/10.1079/PAVSNNR201914025

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T. et al. (2016) Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, **186**, 64–87. https://doi.org/10.1016/J.RSE.2016.08.013

Fung, T. & LeDrew, E. (1988) The determination of optimal threshold levels for change detection using various accuracy indices. *Photogrammetric Engineering & Remote Sensing*, **54**, 1449–1454.

Hall, R.J., Castilla, G., White, J.C., Cooke, B.J. & Skakun, R.S. (2016) Remote sensing of forest pest damage: a review and lessons learned from a Canadian perspective. *Canadian Entomologist*, **148**, S296–S356. https://doi.org/10.4039/tce.2016.11

Henrich, V., Götze, C., Jung, A., Sandow, C., Thürkow, D. & Cornelia, G. (2009) *Development of an online indices database: motivation, concept and implementation*.

Hill, R.A., Wilson, A.K., George, M. & Hinsley, S.A. (2010) Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science*, **13**, 86–99. https://doi.org/10.1111/j.1654-109X.2009.01053.x

Horler, D.N.H., Dockray, M. & Barber, J. (1983) The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, **4**, 273–288. https://doi.org/10.1080/01431168308948546

Housman, I., Chastain, R. & Finco, M. (2018) An evaluation of forest health insect and disease survey data and satellite-based remote sensing forest change detection methods: case studies in the United States. *Remote Sensing*, **10**, 1184. https://doi.org/10.3390/rs10081184

Karnosky, D.F. (1979) Dutch Elm Disease: a review of the history, environmental implications, control, and research needs. *Environmental Conservation*, **6**, 311–322. https://doi.org/10.1017/S037689290000357X

Kemsley, E.K. (1996) Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems*, **33**, 47–61. https://doi.org/10.1016/0169-7439(95)00090-9

Kher, A., Mulholland, M., Green, E. & Reedy, B. (2006) Forensic classification of ballpoint pen inks using high performance liquid chromatography and infrared spectroscopy with principal components analysis and linear discriminant analysis. *Vibrational Spectroscopy*, **40**, 270–277. https://doi.org/10.1016/J.VIBSPEC.2005.11.002

Knipling, E.B. (1970) Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*, **1**, 155–159. https://doi.org/10.1016/S0034-4257(70)80021-9

Landis, J.R. & Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159. https://doi.org/10.2307/2529310

Lee, L.C., Liong, C.-Y. & Jemain, A.A. (2018) Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst*, **143**, 3526–3539. https://doi.org/10.1039/C8AN00599K

Littlewood, N.A., Nau, B.S., Pozsgai, G., Stockan, J.A., Stubbs, A. & Young, M.R. (2015) Invertebrate species at risk from Ash Dieback in the UK. *Journal of Insect Conservation*, **19**, 75–85. https://doi.org/10.1007/s10841-014-9745-2

Lu, D. & Weng, Q. (2007) A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, **28**(5), 823–870. https://doi.org/10.1080/01431160600746456

Mansfield, J.W., Galambos, N. & Saville, R. (2018) The use of ascospores of the dieback fungus *Hymenoscyphus fraxineus* for infection assays reveals a significant period of biotrophic interaction in penetrated ash cells. *Plant Pathology*, **67**, 1354–1361. https://doi.org/10.1111/ppa.12844

McKinney, L.V., Nielsen, L.R., Collinge, D.B., Thomsen, I.M., Hansen, J.K. & Kjaer, E.D. (2014) The ash dieback crisis: genetic variation in resistance can prove a long-term solution. *Plant Pathology*, **63**, 485–499. https://doi.org/10.1111/ppa.12196

Murfitt, J., He, Y., Yang, J., Mui, A. & De Mille, K. (2016) Ash Decline assessment in emerald ash borer infested natural forests using high spatial resolution images. *Remote Sensing*, **8**, 256. https://doi.org/10.3390/rs8030256

Naes, T. & Mevik, B.-H. (2001) Understanding the collinearity problem in regression and discriminant analysis. *Journal of Chemometrics*, **15**, 413–426. https://doi.org/10.1002/cem.676

Pautasso, M., Aas, G., Queloz, V. & Holdenrieder, O. (2013) European ash (*Fraxinus excelsior*) dieback – a conservation biology challenge. *Biological Conservation*, **158**, 37–49. https://doi.org/10.1016/J.BIOCON.2012.08.026

Pietrzykowski, E., Sims, N., Stone, C., Pinkard, L. & Mohammed, C. (2007) Predicting *Mycosphaerella* leaf disease severity in a *Eucalyptus globulus* plantation using digital multi-spectral imagery. *Southern Hemisphere Forestry Journal*, **69**, 175–182. https://doi.org/10.2989/SHFJ.2007.69.3.7.357

Pontius, J., Martin, M., Plourde, L. & Hallett, R. (2008) Ash decline assessment in emerald ash borer-infested regions: a test of tree-level, hyperspectral technologies. *Remote Sensing of Environment*, **112**, 2665–2676. https://doi.org/10.1016/J.RSE.2007.12.011

R core team. (2020) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rackham, O. (1975) *Hayley Wood, its history and ecology*. 1976. Trees Woodl. Br. Landsc.

Rackham, O. & Coombe, D.E. (1996) Madingley Wood. *Nature in Cambridgesh*, **38**, 27–54.

Reinmann, A.B. & Hutyra, L.R. (2017) Edge effects enhance carbon uptake and its vulnerability to climate change in temperate broadleaf forests. *Proceedings of the National Academy of Sciences*, **114**, 107–112. https://doi.org/10.1073/pnas.1612369114

Stone, C. & Mohammed, C. (2017) Application of remote sensing technologies for assessing planted forests damaged by insect pests and fungal pathogens: a review. *Current Forestry Reports*, **3**, 75–92. https://doi.org/10.1007/s40725-017-0056-1

Toloşi, L. & Lengauer, T. (2011) Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics*, **27**, 1986–1994. https://doi.org/10.1093/bioinformatics/btr300

Waser, L., Küchler, M., Jütte, K. & Stampfer, T. (2014) Evaluating the potential of worldview-2 data to classify tree species and different levels of ash mortality. *Remote Sensing*, **6**, 4515–4545. https://doi.org/10.3390/rs6054515

Weaving, D., Jones, B., Ireton, M., Whitehead, S., Till, K. & Beggs, C.B. (2019) Overcoming the problem of multicollinearity in sports performance data: a novel application of partial least squares correlation analysis. *PLoS One*, **14**, e0211776. https://doi.org/10.1371/journal.pone.0211776

Wu, C. (2004) Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment*, **93**, 480–492. https://doi.org/10.1016/J.RSE.2004.08.003

Yadav, S. & Shukla, S. (2016). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In: *Proceedings - 6th International Advanced Computing Conference. IACC 2016*. Bhimavaram: Institute of Electrical and Electronics Engineers Inc., pp. 78–83. https://doi.org/10.1109/IACC.2016.25

Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernández-Clemente, R. et al. (2018) Previsual symptoms of Xylella fastidiosa infection revealed in spectral plant-trait alterations. *Nature Plants*, **4**, 432–439. https://doi.org/10.1038/s41477-018-0189-7

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Summary of number of sampled trees per species and disease class in the four forest study sites near Cambridge.

Table S2. The names and formulae of vegetation indices (VIs) calculated from the hyperspectral data.

Figure S1. The overall accuracies of the PLS-DA species classification models plotted against the number of crowns per species used to create the model.

Figure S2. The reflected spectra of tree species and grass (A) along with the associated coefficient of variation (CV) of the reflectances across the spectrum (B).

Figure S3. The first two principal components (PCs) produced by applying principal components analysis (PCA) on the species training dataset.

Figure S4. Scatter plot of the second latent variable (LV2) against the first latent variable (LV1) obtained from PLS-DA.

