Index measures for oak decline severity using phenotypic descriptors

Jasen P. Finch [a,*], Nathan Brown [b], Manfred Beckmann [a], Sandra Denman [b], John Draper [a]

[a] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK
[b] Centre for Forestry and Climate Change, Forest Research, Alice Holt Lodge, Farnham GU 4LH, UK

ABSTRACT

Oak decline syndromes such as Acute Oak Decline (AOD) and Chronic Oak Decline (COD) are becoming increasingly prevalent and with this comes the need for more quantitative, sensitive and standardised visual oak health monitoring.

Phenotyping protocols were developed to specifically measure oak decline severity and were based on a comprehensive set of simple to measure phenotypic descriptors. A total of 36 phenotypic measurements describing oak decline status included aspects of tree size, crown condition, the presence of biotic agents and a number of derived composite descriptors.

Phenotypic measurements were collected from a total of 174 *Quercus robur*, surveyed from 9 sites across England and included healthy, AOD, COD and AOD trees in remission. Using these data, the Phenotypic Decline Index (PDI) and the Decline Acuteness Index (DAI) were developed to quantitatively describe and differentiate the acute and chronic oak decline severity spectrum. These decline indexes were derived from unsupervised random forest machine learning models, trained using the collected phenotypic data.

The suitability of the decline indexes for describing decline severity and type were assessed by comparing decline index scores to manual decline status classifications along with an assessment of descriptor importance and contribution to the decline index models. Crown condition and trees size descriptors such as 'composite crown volume' contributed positively to the PDI. Trees with smaller crowns in poor condition had greater PDI values. Tree stature and the presence of stem bleeding contributed highly to the DAI, allowing differentiation between trees with AOD and COD syndromes. AOD trees had relatively larger stature and the presence of stem bleeding where as COD trees had small stature and stem bleeding was absent.

The oak decline indexes are simple but sensitive measures of tree decline severity and allow easy comparisons of oak trees both spatially and temporally. These have the potential to provide useful tools for forest monitoring and management as well application to remote sensing and omics research.

1. Introduction

Trees and forests across the globe face threats from a changing climate. Notably warming and an increase in drought episodes can adversely affect the health status of individual trees, causing vegetation changes at the landscape scale (Choat et al., 2012; Millar and Stephenson, 2015; Allen et al., 2015). Within this context, biotic agents (pests and diseases) can flourish and often drive mortality (Kautz et al., 2017; Jactel et al., 2012). Monitoring vegetation changes and understanding the underlying processes affecting tree health are important first steps to predicting the scale and extent of these threats (Hartmann et al., 2018) and plan future policies for planting and mitigation efforts (Bradford and Bell, 2017). Current omics and remote sensing

approaches provide powerful tools to document and understand these changes across multiple scales (from cellular to tree to landscape scales), but an important prerequisite is the ability to accurately describe tree health status, so that even subtle changes can be documented (Dungey et al., 2018; Hornero et al., 2020). Phenotyping, documenting the physical traits of an individual or group, is thus a crucial first step relevant to studies across disciplines and further, it is necessary to understand how these measurements combine to describe health status.

Oak decline is expected to become increasingly prevalent in the UK and across Europe with the onset of climate change (Sturrock et al., 2011; Petr et al., 2014; Brown et al., 2018). Oak declines are complex syndromes that are facilitated by a variety of both abiotic and biotic factors. Abiotic factors can include both long-term predisposing and

one-off inciting factors. Biotic agents can include insect infestation such as larvae of the bark boring beetle *Agrilus biguttatus* or bacterial stem bleeds associated with Acute Oak Decline (AOD) (Brown et al., 2017b; Denman et al., 2014), as well as a variety of fungal root pathogens associated with Chronic Oak Decline (COD) (Denman et al., 2017). The bacterial stem bleeds associated with AOD initially start as active dark weeping patches. Over time, as lesion development continues beneath the bark, external cracks form between the bark plates that are surrounded by blackening stains. Evidence of stem tissue repair can be seen as lignified-callus-like tissue that, when extensive, can fully occlude the internal lesion, appearing as exposed rolls of callused tissue on the exterior (Denman et al., 2014). Stem bleeds can also be associated with a range of root pathogens including *Armillaria* attack on the inner stem tissues of oak. Active stem weeping and black staining can be found along the leading edge of the mycelial front of affected stem tissues (Denman et al., 2017).

Oak decline syndromes are wide ranging in cause and attempts have been made to subdivide the types of decline associated with *Quercus robur* and *Q. petraea* based on the duration over which a tree has been declining and the biotic agents associated with a declining tree. Mortality in trees suffering from COD can occur over a period of decades whereas AOD can occur in 5–10 years (Denman and Webber, 2009). A chronically declining oak can also subsequently begin to develop the stem bleeds associated with AOD (Denman et al., 2017).

As a result of the variety of factors contributing to oak decline and oak responses to these, there are many visual phenotypic descriptors that contribute to categorizing an oak tree as being in decline. When the decline is severe there are clear symptoms visible in the crown, such as low foliage density and numerous dead branches. It is these characteristics that many studies use as the primary indicator of tree decline severity (Zarnoch et al., 2004), however symptoms can develop slowly (especially in COD) and simple classes are unlikely to capture the full detail of the subtle transition from being healthy to declined.

The visual, manual assignment of decline status and severity in the field is variable and error prone due to the many factors contributing to oak decline and surveyor bias. For instance, accurate and reproducible crown condition assessments often require rigorous quality assurance steps that include training and regular calibration among surveyors (Ferretti, 1997; Redfern and Boswell, 2004). Additionally, differences in woodland age and individual tree sizes at different geographical locations can add difficulty to objectively comparing health status between sites. These factors require standardisation to reduce the potential confounding comparisons for tree health status (Zarnoch et al., 2004).

Assigning discrete decline status classes neglects the true spectrum of decline severity encountered in the field. There is a growing need for sensitive, continuous measures of tree decline that accurately reflect the decline severity spectrum (Pontius and Hallett, 2014). This has the potential to greatly improve monitoring, enable the early detection of decline for accurate forecasting and inform forestry management practices. Consistent long term monitoring has been highlighted as a key requirement for understanding the threats to oak and safeguarding its future (Hartmann et al., 2018). Coupled with this, monitoring methods should take advantage of modern data analysis techniques whilst still retaining practical simplicity in the field.

Machine learning is finding increasing utility in ecological and environmental applications due to it's ability to learn complex patterns within and between variables (Cutler et al., 2007; Belgiu and Drăguţ, 2016). A popular algorithm is random forest due to it's versatility, relatively simple model training requirements and ease of use. It is an ensemble machine learning method where forests of bootstrap sampled decision trees are grown, yielding both 'out of bag' sample prediction accuracy and variable importance (Breiman, 2001a; Biau and Scornet, 2016).

Here, we propose a framework using unsupervised random forest machine learning to enable quantitative, visible oak decline severity assessments that utilise a range of oak decline informative phenotypic

descriptors, observed in oak forests across England.

The two resulting measures are:

- The Phenotypic Decline Index (PDI) - a measure of decline severity, scoring trees between 0 and 1. More severely declining oak trees have a score closer to 1.
- The Decline Acuteness Index (DAI) - a measure differentiating between chronically and acutely declining oak trees with a score between −1 and 1. Acutely declining trees have a score closer to 1 and chronically declining trees have a score closer to −1.

2. Materials and methods

2.1. Phenotypic data collection

The phenotypic data were collected over two summers between June and September from a total of 174 trees at nine sites across the south of England (Table 1). The occurrence of AOD has previously been reported and studied at these sites (Brown et al., 2016; Brown et al., 2017b). These sites included a variety of woodlands with diverse management histories, tree sizes/ages and were made up of mainly *Q. robur*.

All trees were manually assigned a decline status category; healthy, AOD, COD or remission prior to collection of phenotypic descriptors. At sites with only COD present, trees were either COD (trees with clear signs of poor crown condition and dieback) or healthy (dominant or co-dominant trees with full crowns). At Site 2, trees with stronger crowns were also categorised as COD when *Armillaria* (*Desarmillaria tabescens*) fruiting bodies had been observed on their buttress roots. At sites with AOD present, annual monitoring data collected since 2009 was used to guide tree selection. Healthy trees had no known history of AOD and fruiting bodies of root pathogens had not been observed; all trees had crown condition of 4 or 5, see Brown et al. (2016) for details of this

Table 1
Overview of the surveyed site information.

Site	UK Region	Site description	Oak decline types present	Survey date	Trees surveyed
1	West Midlands	Oak dominated amenity woodland, parkland plantation.	AOD/COD	June 2017	21
2	East of England	Oak dominated woodland managed for timber.	AOD/COD	September 2016	22
3	South West	Oak dominated plantation managed for timber.	COD	August 2016	20
4	East of England	Oak dominated woodland managed for timber, with lapsed lime coppice.	AOD/COD	June 2017	20
5	South East	Parkland with shelterbelts.	AOD/COD	July 2017	20
6	South West	Open high forest, mown grass below.	AOD/COD	June 2017	22
7	London	Oak dominated urban woodland, boundary belt of park.	AOD/COD	August 2017	9
8	South West	Oak dominated amenity woodland.	COD	August 2016	20
9	East of England	Oak dominated woodland managed for timber.	AOD/COD	June 2017	20

assessment. AOD trees had current stem bleeds but a variety of crown conditions. COD trees either had poor crowns (category 2) or the fruiting bodies of *Gymnopus* or *Armillaria* species had been observed in previous years. The remission group included trees that had shown stem bleeds associated with AOD within the previous three years but were not showing signs of stem bleeds at the time of assessment. Thus, at all sites, selection attempted to create a balanced sample of condition categories including both the unhealthiest and healthiest trees.

The phenotypic descriptors related to tree size, crown condition and the presence of biotic agents including insects, bacteria and fungi are listed and described in Supplementary Table S1. Surveyors involved in phenotyping the trees used in this study had prior experience conducting ICP-Forests level I crown condition surveys. In each year of sampling, surveyors underwent refresher training and "calibration" assessments in oak woodland. The descriptors collected were based on Innes (1990) and ICP-Forests (2016). Diameter at breast height (DBH) was collected at a stem height of 1.3 m, to the nearest 1 mm. Total height, timber height and lowest live crown were measured using a Vertex IV Digital Hypsometer (Haglof, Sweden) to the nearest 0.1 m. The timber height was measured at either where the upper stem diameter was 70 mm or where no main stem could be distinguished. For crown contact, missing crown and absolute crown transparency, percentages were visually estimated between 0 and 100% in 5% interval ranges.

Other categorical descriptors included social class, dieback type, main crown dieback location, stem epicormic growth, branch epicormic growth, extent of insect damage, insect damage type, extent of crown mildew and the presence or absence of pruning damage or branch loss from the main stem. See Supplementary Table S1 for category details.

Crown radius, canopy closure and bark tap tests for dead stem tissue were measured for each cardinal point. The crown radius was measured using a Vertex IV Digital Hypsometer attached to the relevant side of the stem at a height of 1.3 m. Canopy closure was recorded as yes or no depending on whether any other crowns were touching the sampled tree. A high resolution photo was taken of the base to the top of the crown and of the stem from the base to a height of 3 m for each cardinal point for later reference.

The AOD stem symptoms (active bleeds, black staining, callused wounds) were collected by measuring longest axis of each individual bleed, stain or callused wound found within the basal 3 m of the stem. The frequency of *A. biguttatus* exit holes was counted along with the presence or absence of other oval shaped insect exit holes, small circular shaped exit holes and ground level fungal fruiting bodies which were recorded around the stem circumference from the basal tree collar to a radial distance of 2 m from the stem.

2.2. Phenotypic descriptor preparation

R version 4.0.3 was used for all parsing, preparation and subsequent analysis of phenotypic data (R Core Team, 2020). All raw data, data preparation routines and analysis code used to generate this manuscript are available at https://github.com/jasenfinch/Index_measures_for_oak_decline_severity_using_phenotypic_descriptors.

All cardinal assessments were averaged to give a single value per tree. The crown density (*CD*) values collected at Site 1 and 6 were converted to crown transparency ($t = 100 - CD$). All descriptors collected in cm were converted to mm except for DBH which was converted to m. The bark tap tests that were positive for hollow stem tissue were converted to present and absent then renamed dead stem tissue. The canopy closure descriptor was converted from yes and no to present and absent. Finally, the stem symptoms lengths for active bleeds, black staining and callused wounds were averaged and a frequency per tree given by the total observed. After preparation there were a total of 31 raw descriptors for analysis.

2.3. Composite descriptor calculation

Five additional, composite descriptors were calculated to account for the three dimensional attributes of the tree and covariate nature of certain descriptors as suggested by Zarnoch et al. (2004). This then gave a total of 36 raw and composite phenotypic descriptors. The data for these phenotypic descriptors for all 174 trees can be found in Supplementary Table S3. The composite descriptors are defined as follows:

Live crown ratio (%) - percentage of the whole tree height that supports live crown.

$$R = \frac{(h - l)}{h} \times 100$$

h = total height (m), l = lower crown height (m).

Crown condition (%) - percentage of present crown proportional to the percentage crown density.

$$c = (100 - m)(1 - \frac{t}{100})$$

m = missing crown (%), t = crown transparency (%)

Composite crown volume (m^3) - estimated volume of the crown as a paraboloid proportional to the crown condition. Adapted from the composite crown volume from Zarnoch et al. (2004) using crown condition instead of crown density alone.

$$v = \frac{1}{2}\pi r^2 (h - l) \times \frac{c}{100}$$

r = crown radius (m).

Estimated bleed prevalence (%) - estimated percentage of the surveyed 3 m basal trunk area affected by bleeds or black staining.

$$p = \frac{a^2 A + b^2 B}{3000 d\pi} \times 100$$

a = average active bleed length (mm), A = number of active bleeds, b = average black stain length (mm), B = black stain number, $d = DBH(m) \times 1000$.

Agrilus exit hole density (m^{-2}) - density of observed *A. biguttatus* exit holes across the surveyed 2 m basal trunk area.

$$e = \frac{h}{2d\pi}$$

h = number of *A. biguttatus* exit holes, d = DBH (m).

2.4. Random forest analysis of inter-site discrimination

Supervised random forest classification was performed to assess the discrimination in the phenotypic descriptors between the site locations and identify descriptors that could potentially confound comparisons relating to oak decline severity. Random forest analysis was performed using the randomForest R package version 4.6.14 (Liaw and Wiener, 2002) with 100 repetitions using default *mtry* and *ntree* parameters. Due to uneven class sizes, the site classes were balanced using stratified sampling with a sample size of 9 per class, the smallest class size.

To assess classification performance the *margin* metric was computed for each observation using the out of bag predictions, which was averaged across all observations for each random forest repetition, and then averaged across all repetitions. The *margin* metric is the difference between the proportion of votes for the correct class and the maximum proportion of votes for the other classes for a given observation. A positive *margin* value indicates correct classification (Breiman, 2001a).

Descriptor importance contributions were measured using the mean decrease in accuracy metric averaged for each descriptor across the repetitions. Higher mean decrease in accuracy values had greater overall contribution and therefore greater importance.

2.5. Tree size descriptor adjustment for age and inter site variance

To account for age, size and site specific tree differences, the tree size descriptors (DBH, lower crown height, timber height, total height, crown radius) were standardised using an adjustment of site means to the overall mean. An example of this adjustment is shown for DBH in Supplementary Fig. S1 and the adjustment factors for each of the descriptors at each site are given in Supplementary Table S2.

All composite descriptors utilising these adjusted descriptors were recalculated after adjustment. The application and effect of this adjustment is described and discussed in Sections 3.1 and 4.3 respectively.

2.6. Calculation of indexes for oak decline severity

The phenotypic descriptors were analysed using 100 repetitions of unsupervised random forest. The proximity matrices from each repetition were averaged to a single proximity matrix (P). This was then transformed to a dissimilarity matrix ($1 - P$) and scaled to two dimensions using multidimensional scaling (MDS). The first MDS dimension (D_1) was inverted ($1 - D_1$) to ensure the correct dimensional orientation relative to decline severity. It was then min–max scaled to give a scale between 0 and 1 with more severely declining trees having a score closer to 1. This dimension was named the Phenotypic Decline Index, PDI. The second MDS dimension (D_2) was min–max scaled and then transformed ($2D_2 - 1$) to give a scale between -1 and 1 with COD trees having a score closer to -1 and AOD having a score closer to 1. This dimension was named the Decline Acuteness Index, DAI. The code used to generate the decline indexes is available as an R package pdi version 0.4.1 (Finch, 2020).

2.7. Descriptor contribution to decline indexes

To investigate the descriptor contributions to the decline indexes, supervised random forest regression models were trained for each decline index using the phenotypic descriptors for all 174 trees. These models will subsequently be referred to as M_{PDI} and M_{DAI} for the PDI and DAI respectively. The parameters number of trees (*ntree*) and number of randomly assigned variables (*mtry*), for each random forest model, were tuned prior to training using the values of 10, 100, 1,000, 10,000 for *ntree* and a range of 2–26 at intervals of 4 for the *mtry* parameter. The combination of parameters giving the lowest mean absolute percentage error (MAPE) were selected and set as follows: for the PDI (M_{PDI}): *ntree* = 1000, *mtry* = 14, MAPE = 4.74%; for the DAI (M_{DAI}): *ntree* = 10000, *mtry* = 14, MAPE = 6.2% (Supplementary Fig. S2). Overall descriptor importance contributions to each of the indexes was assessed using the percentage increase in mean squared error measure with higher values indicating a greater overall contribution to the given model and therefore greater importance.

2.8. Local descriptor contribution to index models

To identify which phenotypic descriptors contributed locally to specific example cases of each of the decline indexes, Local Interpretable Model-Agnostic Explanations (LIME) were computed for each of models M_{PDI} and M_{DAI} (Ribeiro et al., 2016). The example cases for the PDI used for M_{PDI} included the trees with the minimum (0) and maximum (1) PDI values and the tree with the PDI value closest to the median (0.522) for the healthy, severe and moderately declining trees respectively. For the DAI (M_{DAI}), five example cases were selected from trees with PDI scores greater than 0.5. These cases included a decline neutral case with a score closest to 0 and then for each of the AOD and COD syndromes, a moderate and a severe case were selected with scores closes to 0.5 and 1 for AOD and -0.5 and -1 for COD respectively. LIME analysis was conducted using the lime R package version 0.5.1 with the default parameters (Pedersen and Benesty, 2019).

2.9. Index model simulations to investigate interactions of key descriptors

The interactions of key raw descriptors identified with high percentage increase in mean squared error importance scores and their associations with the decline indexes was investigated by using simulated descriptor values to construct index model (M_{PDI}, M_{DAI}) response surfaces for specific example cases. The same example cases were used as described for the LIME analyses in Section 2.8. This approach was similar to response surface methodology without the need for probabilistic approximation due to the low computation time required for index calculation (Abbasi and Mahlooji, 2012).

For M_{PDI}, missing crown and crown transparency were compared for example cases of a healthy, a moderate and a severely declining tree. The value ranges were based on the observed range in the training data. For the missing crown (0–95%) and crown transparency (0–95%) descriptors, the ranges were divided into 100 increments. For a given case, each combination of these simulated descriptor values were used to calculate a PDI value from M_{PDI}, building a model response surface.

The same approach was used for M_{DAI} using a comparison between active bleed length and black staining length as well as total height and lower crown height. The upper boundary for the ranges of active bleed length and black stain length was taken as the centre value between the maximum and minimum values in the training data sets. These were reduced by a factor of eight to give ranges of 0–8.75 mm and 0–8.12 mm respectively and enabled more effective visualisation. For the total height (16.2–31.1 m) and lower crown height (4–16.2 m) descriptors, the ranges were divided into 100 increments.

3. Results

3.1. Phenotypic differences between site locations and trees size descriptor adjustments

A preliminary analysis using supervised random forest classification could separate trees based on their location (Supplementary Fig. S3a) with an average *margin* value of 0.231. The descriptors that contributed most to this discrimination were those that were related to tree size such as DBH, total height and crown radius and likely reflect the varied tree age and planting density across the nine sites (Supplementary Fig. S3b). To counter this effect, sizes were adjusted based on the site means for each tree size descriptor to remove inter-site variability but maintain intra-site variability as described in Section 2.5. This ensures that the subsequent comparisons of these descriptors in the context of oak decline are not confounded by these identified site differences. After the application of the tree size adjustment, the average random forest *margin* value was substantially reduced to 0.0727. The reduction in this value indicates that the adjustment was able to remove the majority of the between site variance and that the random forest models were no longer able to accurately classify the site locations of the trees (Supplementary Fig. S4a) with almost no contribution of the corrected tree size descriptors (Supplementary Fig. S4b).

3.2. Oak decline severity indexes of manually assigned decline statuses

The calculated PDI and DAI scores (see Supplementary Table S4) derived from MDS of proximity values from unsupervised random forest for all 174 trees, for which phenotypic descriptor data was collected, is shown in Fig. 1. The healthy trees grouped towards the lower end of the PDI scale and the symptomatic COD and AOD groups towards the higher end. There was a significant difference when the PDI values of the healthy (non-symptomatic) and symptomatic (AOD and COD) decline statuses were compared using an independent two sample t-test (t(144) = 12.4, p < 0.001) with group means of 0.297 and 0.696 for non-symptomatic and symptomatic statuses respectively.

On the DAI scale, the AOD trees grouped at the upper end of the scale (>0) and the COD trees at the lower end (< 0). This was confirmed with
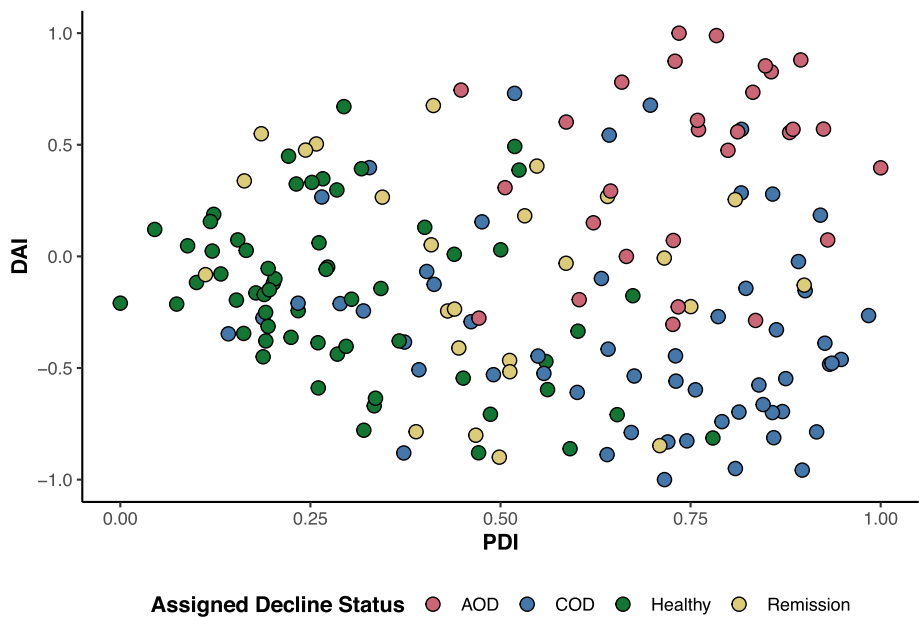
Fig. 1. Decline indexes for all 174 surveyed trees. The decline status categories were manually assigned in the field prior to data collection. AOD = acute oak decline; COD = chronic oak decline. Remisson tress were those that were known to have had active bleeds within the previous 3 years.
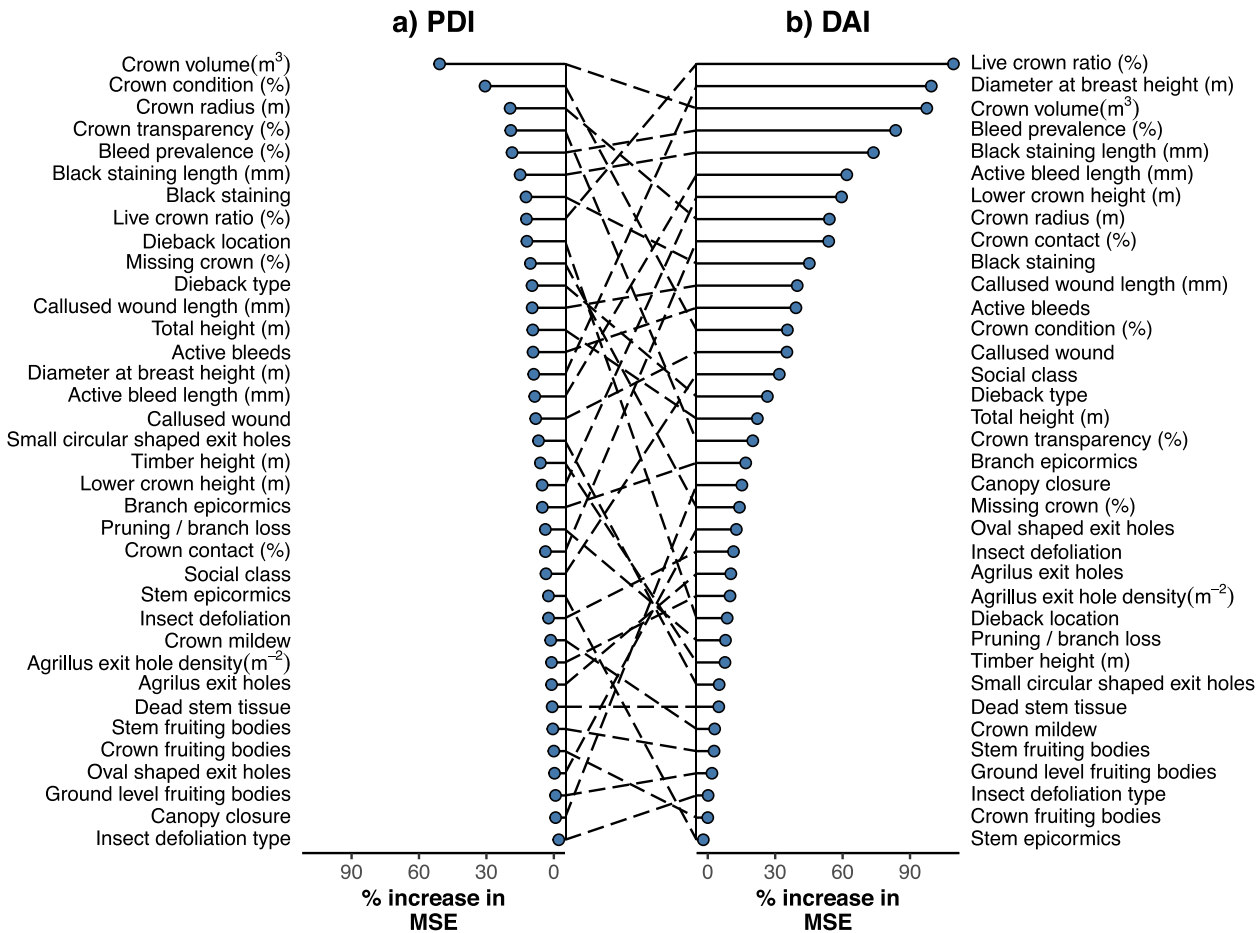

Fig. 2. Descriptor importance scores from the random forest models M_{PDI} and M_{DAI} for relative contributions to the decline index measures. Descriptors are ordered by descending percentage increase in mean squared error (MSE) and therefore from highest contribution to lowest contribution to the given decline index.

significance found in the DAI between these groups (t(84) = 7.89, p < 0.001) with group means of 0.421 and −0.345 for the AOD and COD groups respectively.

The remission trees were found to be significantly different in their PDI scores from both the healthy (t(83) = 4.28, p < 0.001) and symptomatic (AOD and COD; t(109) = 4.63, p < 0.001) groups with means of 0.297, 0.48 and 0.696 for the healthy, remission and symptomatic groups respectively. As a group, these trees are showing moderate overall decline.

The remission trees were also found to be significantly different to the AOD (t(52) = 4.08, p = < 0.001) and COD (t(80) = 2.6, p = 0.011) groups on the DAI scale with group means of 0.421, −0.0685 and −0.345 for the AOD, remission and COD groups respectively. As a group, the remission trees were found to be neither distinctively AOD or COD.

3.3. Key descriptors contributing to the decline indexes

The relative contributions of the phenotypic descriptors to the decline index models (M_{PDI}, M_{DAI}) is shown in Fig. 2.

The top four ranked descriptors found to contribute to M_{PDI} were all crown size and condition related descriptors with composite crown volume the top ranked descriptor overall (Fig. 2a). Other tree size related descriptors such as bleed prevalence, live crown ratio and stem bleeding descriptors contributed secondarily.

M_{DAI} was more complex with many more descriptors contributing more equally than for the PDI (Fig. 2b). The overall tree size descriptors such as live crown ratio, DBH and crown volume contributed almost equally to M_{DAI}. These were followed closely by the stem staining and bleeding descriptors such as bleed prevalence. The presence of *A. biguttatus* exit holes had very little contribution to the DAI with exit hole density ranked only 25th.

Plots of the individual descriptors against the decline indexes are available in Supplementary Figs. S5-S8. A web application is also available at https://jasenfinch.shinyapps.io/decliner/ that allows interactive input of phenotypic descriptor values into the M_{PDI} and M_{DAI} models. The following Sections 3.4 and 3.5 identify how the descriptors contribute and interact within the decline index models in specific decline cases.

3.4. Local descriptor contributions and trends within the PDI and DAI models

LIME identified crown volume, condition, transparency and radius as the top four descriptors contributing to M_{PDI} in each of the healthy, moderate and severe decline cases (Table 2). These descriptors were also among the same top descriptors in the overall model importance lists for the PDI (Fig. 2a). Increasing decline severity was associated with a reduction in the size and condition of the crown. For instance, declining trees were found to have a composite crown volume below 841 m^3. Healthy trees were found to have a crown condition greater than 71.3%, moderately declining trees between 47.6% and 71.3% and severely declining trees a condition below 23.9%.

For M_{DAI}, a number of different descriptors were found to be locally important to the decline type cases compared to the overall model importance such as crown contact (Table 3). This likely reflects the greater number of descriptors contributing to the DAI with different descriptors relevant to particular decline contexts. The important descriptors identified were mainly associated with tree size such as live crown ratio and lower crown height. Bleed prevalence and stem symptom descriptors were not found to be locally important to any of the decline type cases. Live crown ratio was greater than 61.6% in the severe AOD and moderate COD cases. COD trees were found to have greater crown contact (greater than 50%) compared to AOD trees.

Table 2

LIME results for oak decline severity cases in the PDI index model M_{PDI}. Top four important descriptors shown. Descriptors ordered by descending absolute weight within each case.

Status	PDI	Predicted PDI	Descriptor	Descriptor (*d*) range	Weight
Healthy	0.000	0.0571	Crown volume (m^3)	841 < d <= 1683	−0.1090
			Crown condition (%)	71.3 < d	−0.0584
			Crown radius (m)	5.75 < d <= 8.63	−0.0239
			Crown transparency (%)	d <= 23.8	−0.0238
Moderate decline	0.524	0.4890	Crown volume (m^3)	d <= 841	0.1120
			Crown condition (%)	71.3 < d	−0.0610
			Crown radius (m)	5.75 < d <= 8.63	−0.0246
			Crown transparency (%)	d <= 23.8	−0.0235
Severe decline	1.000	0.9600	Crown volume (m^3)	d <= 841	0.1140
			Crown condition (%)	d <= 23.9	0.0722
			Crown radius (m)	2.88 < d <= 5.75	0.0314
			Crown transparency (%)	71.2 < d	0.0224

3.5. Key descriptor interactions within the decline index models

The complex descriptor interactions within the index models (M_{PDI}, M_{DAI}) were investigated by simulating phenotypic data and generating the index values from the supervised index random forest regression models. The index values, generated by the models from these simulated phenotypic data, were then visualised as model response surface plots, to allow interpretation of descriptor interactions between the indexes. Figs. 3 and Fig. 4 show the simulated index value response surfaces under a variety of different circumstances for the key phenotypic descriptors described in Section 3.3 that contribute highly to the PDI and DAI.

Crown condition descriptors such as composite crown volume and crown condition contributed highly to M_{PDI} (Fig. 2a). The phenotypic measures central to this were percentage missing crown and percentage crown transparency from which crown condition and crown volume are calculated. It can be seen in Fig. 3 that as the percentage missing crown and percentage crown transparency increase and therefore crown condition and composite crown volume decrease, the PDI increases. A much greater range of PDI values can be obtained from trees with relatively greater stature (e.g. greater crown radius, total height, lower crown height and DBH; Fig. 3a & b), while relatively small stature trees have PDIs of 0.6 or greater even with 100% crown condition.

The DAI (M_{DAI}) is a more complex index with a greater number of descriptors contributing to it than the PDI such as live crown ratio, bleed prevalence, and DBH. Simulated response surfaces for the DAI are shown in Fig. 4. These response surfaces show circumstances under which bleed prevalence related descriptors (Fig. 4a, b & c) and live crown ratio descriptors (Fig. 4d, e & f) change for neutral decline type (trees neither distinctively AOD or COD; Fig. 4a & d respectively), moderate AOD or COD (Fig. 4b & e respectively) and severe AOD or COD (Fig. 4c & f respectively). Increasing active bleeds, black staining length and frequency contributed to greater positive values of the DAI and more severe AOD. Higher DAI values could also be obtained from trees with greater stature (Fig. 4c). Total height and lower crown height have no influence

Table 3

LIME results of oak decline type and severity in the DAI index model M_{DAI}. Top four important descriptors shown. Descriptors ordered by descending absolute weight within each case.

Syndrome	Status	DAI	Predicted DAI	Descriptor	Descriptor (d) range	Weight
	Neutral	−0.000249	−0.0405	Live crown ratio (%)	$61.6 < d$	0.0683
				Crown contact (%)	$d <= 25$	0.0404
				Lower crown height (m)	$7.05 < d <= 10.09$	0.0359
				Branch epicormics	$d = 0$	−0.0242
AOD	Moderate	0.504000	0.4820	Live crown ratio (%)	$41.3 < d <= 61.6$	−0.0589
				Lower crown height (m)	$10.09 < d <= 13.14$	−0.0356
				Crown contact (%)	$75 < d$	−0.0395
				Crown radius (m)	$5.75 < d <= 8.63$	0.0171
	Severe	1.000000	0.8630	Live crown ratio (%)	$61.6 < d$	0.0708
				Crown contact (%)	$d <= 25$	0.0419
				Lower crown height (m)	$7.05 < d <= 10.09$	0.0348
				Diameter at breast height (m)	$0.809 < d <= 1.071$	0.0344
COD	Moderate	−0.508000	−0.4670	Live crown ratio (%)	$61.6 < d$	0.0732
				Lower crown height (m)	$d <= 7.05$	0.0276
				Crown contact (%)	$75 < d$	−0.0357
				Branch epicormics	$d = 0$	−0.0239
	Severe	−1.000000	−0.9000	Live crown ratio (%)	$41.3 < d <= 61.6$	−0.0543
				Lower crown height (m)	$7.05 < d <= 10.09$	0.0341
				Crown contact (%)	$50 < d <= 75$	−0.0302
				Crown radius (m)	$2.88 < d <= 5.75$	−0.0177

on the DAI for trees of neutral decline type (Fig. 4d). Similarly to the PDI, trees with smaller stature and poorer crown condition give lower DAI values (Fig. 4e & f) and therefore more severe COD; however, only narrow ranges of DAI values could be obtained for each case when varying the ranges of these descriptors.

4. Discussion

4.1. Suitability of the decline indexes for describing decline severity and type

The decline indexes presented here provide simple intuitive scores based on a wide number of factors known to be heavily influenced by oak decline severity. These two indexes can be represented as a two-dimensional oak decline continuum under which the spectrum of decline severity and type can be assessed (Fig. 5). The index thresholds for category definitions (healthy, declining, severe decline etc.) shown currently serve only as a guide to aid human interpretation as opposed to definitive thresholds of disease. Further work is needed that includes phenotypic data with greater spatial and temporal distribution to establish these definitive thresholds.

Although significant differences were obtained between the decline status groups on both the PDI and DAI scales, Fig. 1 highlights the inaccuracies in the manual decline status classifications when viewed in an individual tree context. It can be seen that a number of the trees, when compared quantitatively using these measures, could in fact be classified with a different decline state. This is likely due, at least in part, to the inclusion of trees in early stages of decline. For example, trees with visible fruiting bodies of root pathogens were manually classed as COD, even if crown health was still relatively good.

Crown condition and volume were shown to be the key contributors to the M_{PDI}. This is to be expected and are indicators of model suitability using the unsupervised approach as crown condition and crown dimensions are well established as the key components of assessing overall tree health and condition (Schomaker et al., 2007). However, previous studies such as Evans et al. (2012) and Pontius and Hallett (2014) have treated these descriptors as linear progressions of oak health that contribute equally to overall tree condition. It has been shown here that in the given visual definitions of oak decline severity; firstly, factors of overall crown condition and size are unlikely to contribute equally (Fig. 2a) and secondly, it is unlikely that these aspects of crown condition contribute linearly (Fig. 3a).

It was shown that tree size, reflected by the live crown ratio descriptor, contributes most to the DAI in discriminating between COD and AOD trees. The relationship of relative tree stature and the acuteness of oak decline has not previously been shown and likely reflects the period of time under which these trees have been afflicted. AOD occurs over a much shorter period compared to COD and therefore has less time in which to inhibit growth of the tree as a whole.

The high contribution of bleed prevalence related descriptors to M_{DAI} is also expected as this is one of the defining features of AOD (Denman et al., 2014). There was no contribution of *A. biguttatus* exit hole occurrence to the DAI and therefore stem bleeding. This could partly be due to the low number of trees on which *A. biguttatus* exit holes were detected (15 out 174) but it has also been shown previously that trees with bleeds do not always have exit holes, attributed to the fact that not all *A. biguttatus* larvae complete their life cycle (Brown et al., 2017b; Vansteenkiste et al., 2004).

There were 4 manually assigned COD trees that were scored with PDI and DAI values greater than 0.5, suggesting that they were in fact AOD trees. These trees were manually assigned COD status due to the presence of *Armillaria* attack and associated bleeds. While the phenotyping protocol used here made no distinction as to the biotic origin of the stem bleeds, it is interesting to note that not only do these trees have the presence of stem bleeds similar to AOD trees, but also share their relatively larger stature compared to the majority of COD trees. This likely reflects a more acute and aggressive nature of the *Armillaria* infection in these trees with deterioration in crown health having occurred over a shorter time period. It therefore has less time in which to impact overall growth compared to the other COD trees. In this sense the DAI scale can be seen to differentiate decline types based less on the actual biotic agents involved in the interaction but more on the time-scales over which the trees themselves have been afflicted.

AOD severity was associated with overall decline severity in trees having both PDI values greater than 0.7 and DAI values greater than 0.5 (Fig. 1). The association of crown condition and AOD severity has previously been shown by Brown et al. (2016).

The remission trees as a group were shown to be intermediate in decline severity and type to the healthy, AOD and COD trees on both the PDI and DAI index scales. However, this was also the most varied of the decline status groups. It is possible that this could reflect the potential for trees to recover from AOD but this is speculative and it is not possible to test this suggestion further with current observations. Using these decline indexes, remission trees could be more objectively defined in
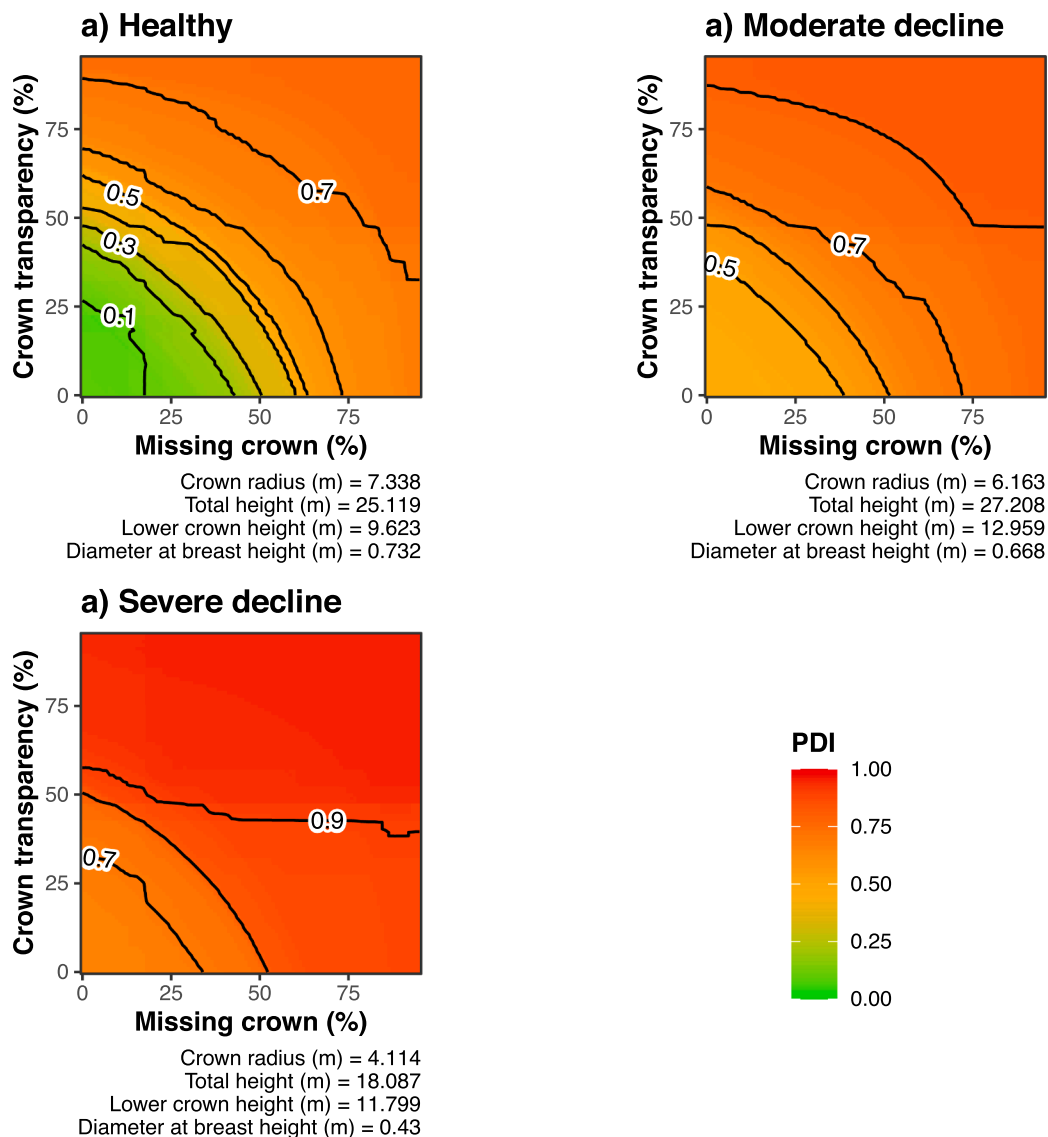
Fig. 3. PDI response surfaces from M_{PDI} for simulated values of key contributing descriptors for example healthy, moderate and severly declining cases. The values of additional tree size descriptors are shown for the individual cases to indicate the relative stature of the example cases.

terms of their temporal trajectory within the decline index continuum.

4.2. The machine learning approach

A machine learning approach was used to derive the phenotypic index measures as it allows the models to be moulded to the complex interactions that exist between the phenotypic descriptors from which the decline indexes are calculated. This would likely be oversimplified using conventional modelling strategies (Breiman, 2001b).

An unsupervised approach was favoured in this context due to the concession that there are known difficulties with and inaccuracy in manual decline status classification in the field (Ferretti, 1997). This is highlighted in Fig. 1 where numerous trees would be considered to have been misclassified as they fall into regions of the decline indexes associated with a different decline status (e.g. COD trees with DAI values >0.5). The unsupervised approach has worked in this context in aptly describing decline severity as the phenotypic descriptors selected were highly specific and well suited to describing the decline severity problem. In situations where descriptors may well be less suited, perhaps as a result of technical limitations, the use of a semi-supervised approach may be favourable. The unsupervised approach to derive the decline

indexes would be applied to a subset of the observations on which the descriptors are known to perform well and that would be considered benchmark observations by the investigators. Indexes could then be attained for the rest of the observations by prediction based on supervised models of the benchmark subset (Chapelle et al., 2010). This approach would be more suitable than a fully supervised approach as it would limit the inaccuracies that would be introduced by manual classification while still allowing the data driven flexibility of the unsupervised approach.

Machine learning approaches are usually seen as black box and hard to interpret; however, there are substantial gains in terms of predictive capability for applying these models in real world situations (Touw et al., 2012). The assessment of descriptor contributions to the index models are an integral part of this approach and helps to overcome these limitations. Knowledge input of the biological mechanisms, by the investigator at this stage, is essential in ensuring that the models suitably learn the patterns relevant to the biological problem to which it is being applied. The descriptors contributing to these should be sensible and it is important to check that the algorithm (random forest) is not learning spurious or artificial patterns within the data such as site differences in tree size and age. Approaches such as LIME and model response surfaces
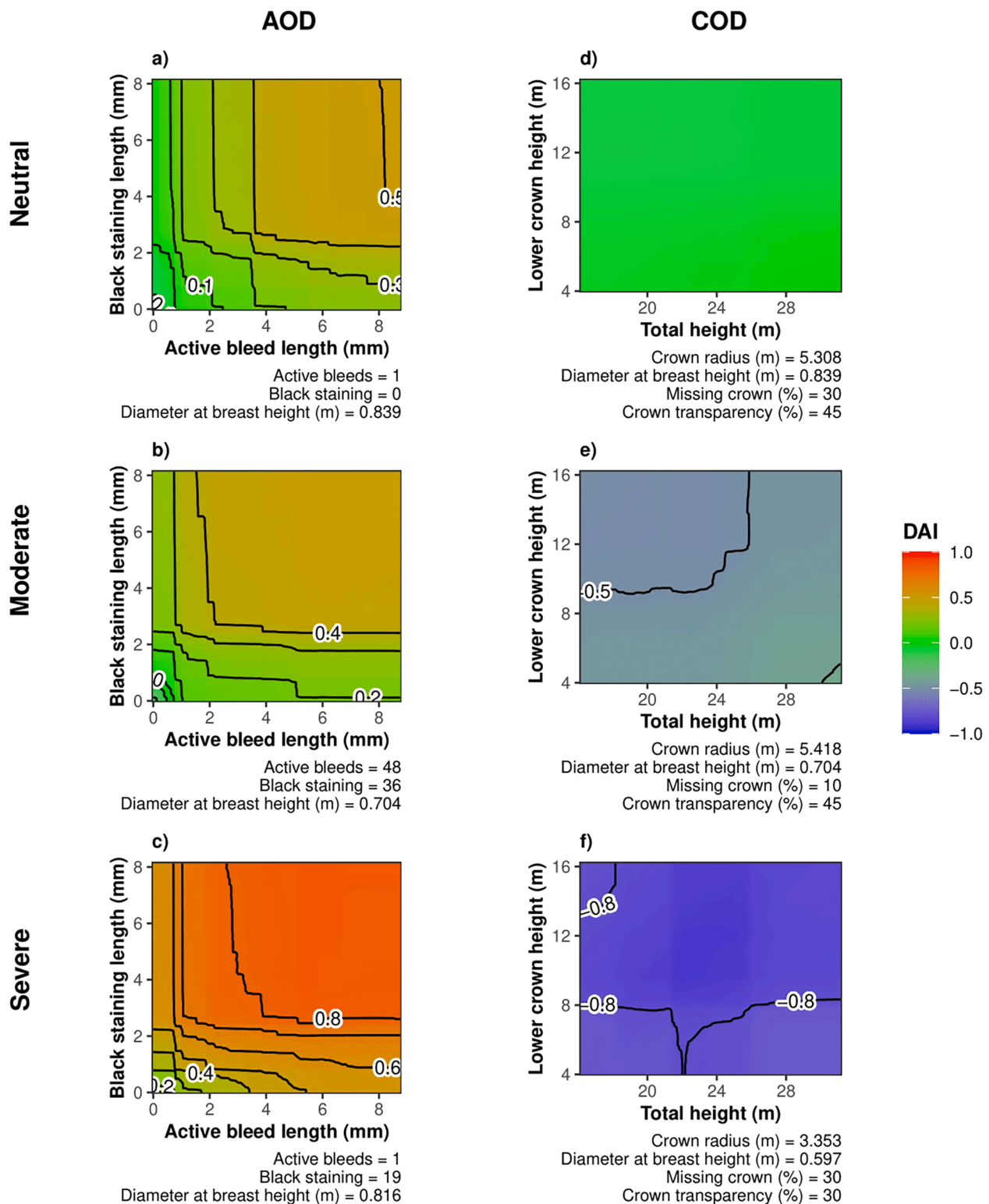
Fig. 4. DAI response surfaces from M_{DAI} for simulated values of key contributing descriptors for example cases. The neutral case is a tree that can be neither distinguished as AOD or COD. All cases have PDI values >0.5. The values of additional descriptors related to bleed prevalence and relative stature/crown condition are shown for the AOD and COD example cases respectively.

enable investigators to query the contributions and interactions of individual phenotypic descriptors. This in turn promotes trust in the indexes derived from these models by the investigator and potential model end-users (Ribeiro et al., 2016).

4.3. Applications

While it is important to note that these decline indexes are study specific and based on a relatively small sample size, they do have the potential for wider use. Subsequent studies could easily build on this foundation by collecting additional data and incorporating it with the
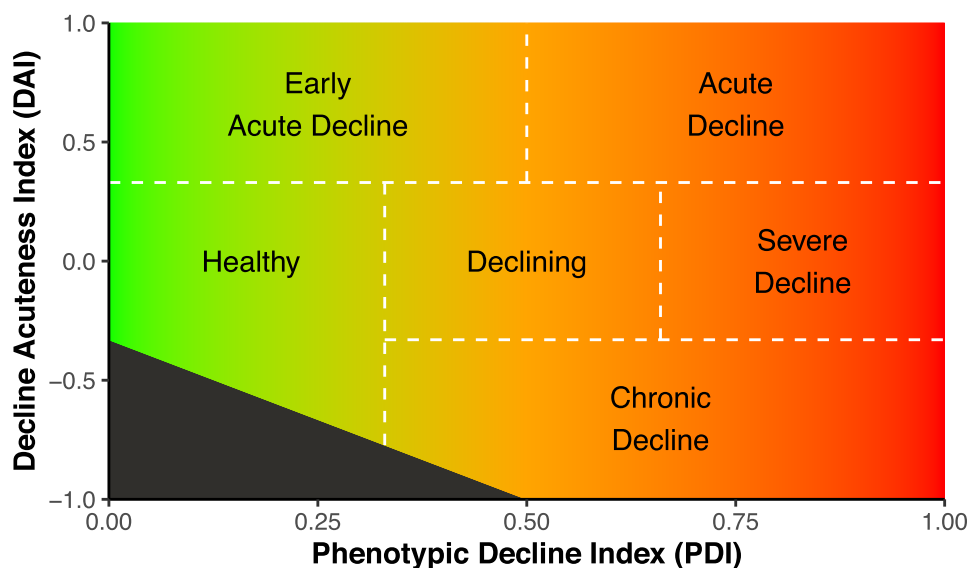
Fig. 5. Two dimensional continuum of the phenotypic decline indexes for oak decline severity. Decline category boundaries intended as a guide to aid interpretation only. Grey area indicates area for which the decline indexes are currently undefined.

data from this study to train random forest models from which to derive decline index values. This would only be viable so long as comparable phenotyping protocols are used. Alternatively, study specific decline index values could be generated using the unsupervised random forest approach.

By defining decline severity objectively across a spectrum, these decline indexes provide greatly enhanced sensitivity over previous categorical decline status assignments. This approach could enhance integration with large scale data such as those from genome wide association studies or for providing ground-truth data for remote sensing (Stocks et al., 2019; Evans et al., 2012). These index measures allow sensitive comparisons of visual health to be made for the same tree between years as well as comparisons of trees between different locations.

Scaling this technique to enable computation of decline indexes for any tree sampled in the UK would require index models trained using a substantially larger data set that incorporates the full descriptive diversity of oak health in the population as whole. Scaling to this level would also require an appropriate method of tree size correction that could be applied on an individual tree basis. A potential solution to this would be to correct trees based on estimated age from DBH as proposed by White (1998). These indexes could have great utility in long term forest health monitoring by providing consistent scores which can be compared over time to inform both forestry management and research (Denman et al., 2010). For instance, the increase of the PDI values of trees within a woodland, beyond what is considered the healthy region of the PDI scale, could indicate that management intervention is required.

The phenotypic descriptors used to calculate these indexes are easily measurable in the field and are based on well established tree condition assessment protocols (Innes, 1990). Further refinement of these protocols could be designed to reduce survey time and complexity and potentially be incorporated into citizen science programmes (Brown et al., 2017a). Alternatively, the machine learning approach used to derive these decline indexes could also be applied to long established oak health monitoring programmes to aid sensitivity in detecting overall trends in tree health. An example programme includes The International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) (Lorenz, 1995).

5. Conclusions

The decline indexes presented here provide a new framework to quantitatively describe the phenotypic oak decline spectrum using intuitive measures of severity (PDI) and type (DAI). Their basis is on simple phenotypic descriptor measurements and a utilisation of a data driven, unsupervised random forest based machine learning approach. This has shown the relative contributions that aspects of tree size, crown condition and symptoms such as stem bleeding have on quantitatively defining our concepts of oak decline status.

Strategies have been presented such as the use of LIME and simulated model response surfaces that can aid the interpretation of the complex descriptor interactions that underpin these simple index measures and greatly reduce the black box nature that is often associated with these approaches. It is encouraged that this approach of deriving descriptive index measures using machine learning is applied in other tree health phenotyping contexts and it is hoped that this will have an impact on the future consistency and comparability of oak health monitoring.

CRediT authorship contribution statement

Jasen P. Finch: Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Nathan Brown:** Methodology, Investigation, Data curation, Writing - review & editing. **Manfred Beckmann:** Writing - review & editing, Supervision. **Sandra Denman:** Methodology, Writing - review & editing, Project administration, Funding acquisition. **John Draper:** Methodology, Supervision, Project administration.

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.

Future Proofing Plant Health (FPPH) Oak Resilience project.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.foreco.2021.118948.

References

Abbasi, B., Mahlooji, H., 2012. Improving response surface methodology by using artificial neural network and simulated annealing. Expert Syst. Appl. 39, 3461–3468.

Allen, C.D., Breshears, D.D., McDowell, N.G., 2015. On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the anthropocene. Ecosphere 6, 1–55.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31.

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227.

Bradford, J.B., Bell, D.M., 2017. A window of opportunity for climate-change adaptation: easing tree mortality by reducing forest basal area. Front. Ecol. Environ. 15, 11–17.

Breiman, L., 2001a. Random forests. Machine Learn. 45, 5–32.

Breiman, L., 2001b. Statistical modeling: The two cultures. Stat. Sci. 16, 199–231.

Brown, N., van den Bosch, F., Parnell, S., Denman, S., 2017a. Integrating regulatory surveys and citizen science to map outbreaks of forest diseases: acute oak decline in england and wales. Proc. Roy. Soc. B: Biol. Sci. 284, 20170547.

Brown, N., Jeger, M., Kirk, S., Williams, D., Xu, X., Pautasso, M., Denman, S., 2017b. Acute oak decline and agrilus biguttatus: the co-occurrence of stem bleeding and d-shaped emergence holes in great britain. Forests 8, 87.

Brown, N., Jeger, M., Kirk, S., Xu, X., Denman, S., 2016. Spatial and temporal patterns in symptom expression within eight woodlands affected by acute oak decline. For. Ecol. Manage. 360, 97–109.

Brown, N., Vanguelova, E., Parnell, S., Broadmeadow, S., Denman, S., 2018. Predisposition of forests to biotic disturbance: Predicting the distribution of acute oak decline using environmental factors. For. Ecol. Manage. 407, 145–154.

Chapelle, O., Scholkopf, B., Zien, A., 2010. Semi-supervised Learning. MIT Press.

Choat, B., Jansen, S., Brodribb, T.J., Cochard, H., Delzon, S., Bhaskar, R., Bucci, S.J., Feild, T.S., Gleason, S.M., Hacke, U.G., et al., 2012. Global convergence in the vulnerability of forests to drought. Nature 491, 752.

Cutler, D.R., Edwards Jr, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Denman, S., Barrett, G., Kirk, S.A., McDonald, J.E., Coetzee, M.P., 2017. Identification of armillaria species on declined oak in britain: implications for oak health. Forestry: Int. J. For. Res. 1–14.

Denman, S., Brown, N., Kirk, S., Jeger, M., Webber, J., 2014. A description of the symptoms of acute oak decline in britain and a comparative review on causes of similar disorders on oak in europe. Forestry: Int. J. For. Res. 87, 535–551.

Denman, S., Kirk, S., Webber, J., et al., 2010. Managing Acute Oak Decline. Forestry Commission.

Denman, S., Webber, J., 2009. Oak declines: new definitions and new episodes in britain. Quart. J. Forestry 103, 285–290.

Dungey, H.S., Dash, J.P., Pont, D., Clinton, P.W., Watt, M.S., Telfer, E.J., 2018. Phenotyping whole forests will help to track genetic performance. Trends Plant Sci.

Evans, B., Lyons, T.J., Barber, P.A., Stone, C., Hardy, G., 2012. Dieback classification modelling using high-resolution digital multispectral imagery and in situ assessments of crown condition. Remote Sens. Lett. 3, 541–550.

Ferretti, M., 1997. Forest health assessment and monitoring–issues for consideration. Environ. Monitor. Assess. 48, 45–72.

Finch, J., 2020. pdi: Phenotypic Index Measures for Oak Decline Severity. https://jasen finch.github.io/pdi. r package version 0.4.1.

Hartmann, H., Moura, C.F., Anderegg, W.R., Ruehr, N.K., Salmon, Y., Allen, C.D., Arndt, S.K., Breshears, D.D., Davi, H., Galbraith, D., et al., 2018. Research frontiers for improving our understanding of drought-induced tree and forest mortality. New Phytol. 218, 15–28.

Hornero, A., Hernández-Clemente, R., North, P., Beck, P., Boscia, D., Navas-Cortes, J., Zarco-Tejada, P., 2020. Monitoring the incidence of xylella fastidiosa infection in olive orchards using ground-based evaluations, airborne imaging spectroscopy and sentinel-2 time series through 3-d radiative transfer modelling. Remote Sens. Environ. 236, 111480.

ICP-Forests, 2016. Manual on methods and Criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Eberswalde, Germany. URL http://icp-forests.net/page/icp-forests-manual.

Innes, J.L., 1990. Assessment of tree condition. Forestry Commission, Great Britain.

Jactel, H., Petit, J., Desprez-Loustau, M.L., Delzon, S., Piou, D., Battisti, A., Koricheva, J., 2012. Drought effects on damage by forest insects and pathogens: a meta-analysis. Glob. Change Biol. 18, 267–276.

Kautz, M., Meddens, A.J., Hall, R.J., Arneth, A., 2017. Biotic disturbances in northern hemisphere forests–a synthesis of recent data, uncertainties and implications for forest monitoring and modelling. Glob. Ecol. Biogeogr. 26, 533–552.

Liaw, A., Wiener, M., 2002. Classification and regression by randomforest. R News 2, 18–22. URL https://CRAN.R-project.org/doc/Rnews/.

Lorenz, M., 1995. International co-operative programme on assessment and monitoring of air pollution effects on forests-icp forests. Water, Air, Soil Pollut. 85, 1221–1226.

Millar, C.I., Stephenson, N.L., 2015. Temperate forest health in an era of emerging megadisturbance. Science 349, 823–826.

Pedersen, T.L., Benesty, M., 2019. lime: Local Interpretable Model-Agnostic Explanations. https://CRAN.R-project.org/package=lime. r package version 0.5.1.

Petr, M., Boerboom, L.G., van der Veen, A., Ray, D., 2014. A spatial and temporal drought risk assessment of three major tree species in britain using probabilistic climate change projections. Climatic Change 124, 791–803.

Pontius, J., Hallett, R., 2014. Comprehensive methods for earlier detection and monitoring of forest decline. Forest Sci. 60, 1156–1163.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Redfern, D., Boswell, R., 2004. Assessment of crown condition in forest trees: comparison of methods, sources of variation and observer bias. For. Ecol. Manage. 188, 149–160.

Ribeiro, M.T., Singh, S., Guestrin, C., 2016. Why should i trust you?: Explaining the predictions of any classifier, in. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1135–1144.

Schomaker, M.E., Zarnoch, S.J., Bechtold, W.A., Latelle, D.J., Burkman, W.G., Cox, S.M., 2007. Crown-condition classification: a guide to data collection and analysis. Gen. Tech. Rep. SRS-102. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station. 78 p. 102.

Stocks, J.J., Metheringham, C.L., Plumb, W., Lee, S.J., Kelly, L.J., Nichols, R.A., Buggs, R.J., 2019. Genomic basis of european ash tree resistance to ash dieback fungus. Nature Ecol. Evol. 1686–1696.

Sturrock, R., Frankel, S., Brown, A., Hennon, P., Kliejunas, J., Lewis, K., Worrall, J., Woods, A., 2011. Climate change and forest diseases. Plant Pathol. 60, 133–149.

Touw, W.G., Bayjanov, J.R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., van Hijum, S.A., 2012. Data mining in the life sciences with random forest: a walk in the park or lost in the jungle? Briefings Bioinformat. 14, 315–326.

Vansteenkiste, D., Tirry, L., Van Acker, J., Stevens, M., 2004. Predispositions and symptoms of agrilus borer attack in declining oak trees. Ann. For. Sci. 61, 815–823.

White, J., 1998. Estimating the age of large and veteran trees in Britain. Forestry Commission, Great Britain.

Zarnoch, S.J., Bechtold, W.A., Stolte, K., 2004. Using crown condition variables as indicators of forest health. Can. J. For. Res. 34, 1057–1070.