
http://researchrepository.murdoch.edu.au/12194/
Enhancing a eucalypt crown condition indicator driven by high spatial and spectral resolution remote sensing imagery

Bradley Evans,a,b Tom Lyons,b Paul Barber,b,c Christine Stone,d and Giles Hardyb

aMacquarie University, Department of Biological Sciences, University Avenue, Macquarie Park, New South Wales, 2113, Australia  
bradley.evans@mq.edu.au  
bMurdoch University, Centre of Excellence for Climate Change, Woodland and Forest Health, 90 South Street, Murdoch, Western Australia, 6150, Australia  
cArbor Carbon Environmental and Arborcultural Consultants P.O. BOX 1065 Willagee Central, Western Australia, 6156, Australia  
dForest Science Centre, NSW Department of Primary Industries, P.O. Box 100, Beecroft, NSW, Australia, 2119

Abstract. Individual crown condition of Eucalyptus gomphocephala was assessed using two classification models to understand changes in forest health through space and time. Using high resolution (0.5 m) digital multispectral imagery, predictor variables were derived from textural and spectral variance of all pixels inside the crown area. The results estimate crown condition as a surrogate for tree health against the total crown health index. Crown condition is derived from combining ground-based crown assessment techniques of density, transparency, dieback, and the regrowth of foliage. This object-based approach summarizes the pixel data into mean crown indices assigned to crown objects which became the carrier of information. Models performed above expectations, with a significant weighted Cohen’s kappa ($\kappa > 0.60$ and $p < 0.001$) using 70% of available data. Using in situ data for model development, crown condition was predicted forwards (2010) and backwards (2007) in time, capturing trends in crown condition and identifying decline in the healthiest between 2008 and 2010. The results confirm that combining spectral and textural information increased model sensitivity to small variations in crown condition. The methodology provides a cost-effective means for monitoring crown condition of this or other eucalypt species in native and plantation forests. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JRS.6.XXXXXX]

Keywords: tree crown condition; forest health; high resolution; digital multispectral imagery; remote sensing; Eucalyptus gomphocephala; tuart; random forest.

1 Introduction

Quantitative indicators of tree crown condition across large areas are a management imperative and critical for the assessment of ecosystem biodiversity, health, and function.1,2 Symptoms of crown decline manifest as a reduction in the abundance, distribution, and greenness of foliage (dieback) within the tree crown and an overall reduction of vigor and condition.3 This dieback can ultimately lead to mortality despite eucalypts being well-adapted to extreme climate variability.4 In recent years, a number of declines in eucalypt dominated ecosystems have been documented.5–8 Similar to the findings in other areas, Eucalyptus gomphocephala (tuart) declines in southwestern Australia have been linked to a range of causes, including soil bacterial communities,9 pathogens,9 and fire frequency and intensity.10–12
Ground-based, or in situ, forest condition assessments of the extent, condition, and distribution of foliage through a tree crown are a common method for assessing changes in health over time. Despite their operational usefulness, they are considered a subjective, semi-quantitative observation and given the size and remoteness of the forest resource, labor intensive and time consuming. As a result, regular monitoring of forest condition over time is difficult yet is a management imperative under changing climatic and land use pressures.

Remote sensing methods and imagery from space-borne or airborne platforms offer an alternative and viable and cost-effective means of assessing forest condition over large areas in a way that is spatially explicit and repeatable through time. Land Monitor II, a Landsat-based vegetation trend product has been demonstrated to be useful in monitoring vegetation changes over time. However, the spatial resolution (30 m) of the Landsat Thematic Mapper (TM) sensor limits the capacity to detect individual tree crowns and in-crown foliage variability. Alternatively, airborne high spatial resolution remote sensing imagery (0.5 m) offers greater flexibility in image acquisition, the capacity to detect individual crowns and crown architecture. When used together, the two resolutions enable the combination of broad area temporal analysis and limited area crown condition analysis.

At both resolutions, the spectral information is often used independently and combined as a vegetation index (VI) to predict condition, such as the distribution and density of foliage. VI summarize spectral information from specific regions of the electromagnetic spectrum known to be sensitive to the interactions between light, leaf, and canopy structure and pigments (mostly chlorophyll). In high resolution data, where many pixels are captured across individual crowns, the spectral characteristics of neighboring pixels was highly variable and as a result, statistics produced in the spatial domain may yield more reliable estimates of crown condition than spectral indicators alone. A number of techniques exist to measure the variation of spectral and spatial properties of remotely sensed images including texture and geometric measures.

Stone and Haywood demonstrated the use of a eucalypt crown condition index derived from remote sensing observations over areas of eucalypt dieback in New South Wales. The approach utilized high spatial resolution digital imagery linked to five key indicators of canopy decline. Recently, it was demonstrated in that a similar index could be used to predict crown dieback in Western Australia from remote sensing imagery. The results indicated that crown condition could be accurately modeled using the spectral properties of the imagery with significant coefficient of determination ($R^2$) values developed from linear regression models ranging from 0.54 to 0.72.

The objectives of this study were threefold. First, we evaluated the additional power of texture and geometric-based indices to predict canopy condition across crowns exhibiting dieback in the Yalgorup National Park of Western Australia. Second, we applied an innovative modeling approach, random forests (RF), a nonparametric, ensemble classifier, to assess the impact of varying the number of predicted canopy condition classes with the aim of developing accurate and robust indicators for forest managers. Lastly, with confidence in the developed models, we then applied them to images acquired before and after the in situ assessments to better understand how the condition of these crowns has changed over time.

2 Methods

2.1 Region of Interest

The Yalgorup National Park, in southwest, Western Australia, is situated 80 vkm south of the capital, Perth (Fig. 1). The sites were located within the Yoongarillup, Karrakatta, and Cottesloe vegetation complexes, based on their geological associations. The majority of the sites were found on the Yoongarillup complex, comprised of E. gomphocephala (tuart), with a mid-storey of Agonis flexuosa, Allocasuarina fraseriana, Banksia grandis, B. attenuata, B. littoralis, and an understorey of Acacia saligna, A. pulchella, Jacksonia sternbergiana, Melaleuca acerosa, and Hibbertia hypericoides. Tuart is listed by the International Union for the Conservation of Nature as IUCN category II species and Yalgorup National Park is an area of high natural biodiversity value.
2.2 Crown Condition Assessment

The placement of crown condition plots within the national park was based on a stratification of vegetation trend classes derived from a 15-year (1990 to 2005) time series analysis of Landsat TM imagery based on the methodology of Caccetta et al. A total of 20 sites were established in 2007 and in June 2008 the crown condition of four randomly selected tuart crowns within each of the 20 sites was individually assessed by two forest health professionals.

Each crown was assessed for four crown condition indices, based on previous research of Barry et al. and recently discussed in Cai et al.; Evans et al.; and Horton et al. The condition indices were measured in 5% classes across a 0% to 100% scale and included (1) crown density (A), (2) foliage transparency (B), (3) crown dieback ratio (C), and (4) an epicormic index (D).

For each crown assessment, two trained forest health professionals stood at right angles to each other at a distance from the main stem greater than its height. Two assessments were assumed to be more robust than a single observation and were averaged for an overall value. The location of each individual crown was mapped using a differential GPS (dGPS) for location in the imagery. All four individual crown condition indices were combined to produce the total crown health index (TCHI), which weights crown density (A), foliage transparency (B), epicormic index (C), and crown dieback ratio (D) equally (i.e., $TCHI = A + B + C + D/4$).

2.3 Image Specification and Acquisition

High spatial resolution airborne digital multispectral imagery (DMSI) was acquired by SpecTerra Services Pty. Ltd. (Perth, Western Australia) in May/June 2007, 2008, and 2010.
under clear conditions (Table 1). During acquisition, all DMSI was georeferenced based on GPS and subsequent post-flight processing of the data included band to band registration and spectro-radiometric correction for bi-directional reflectance distribution function (BRDF) variations. All imagery was acquired at 0.5 m resolution from an altitude of 2000 m. Once collected, all the imagery was mosaicked, orthorectified, and pixel matched to correct for the sun angle differences of each overpass (Table 1).

The DMSI sensor acquires 12-bit digital number (DN) data simultaneously in four narrow bands (20 nm full width half maximum) spectral channels in the visible and near-infrared regions of the electromagnetic spectrum using filters centered at 450, 550, 675, and 780 nm covering blue, green, red and near infrared regions of the spectrum, respectively.

To ensure the relative radiometric consistency of the imagery over the three time periods a “like-value” calibration was undertaken using methods described by Furby et al.35 The method normalizes each scene to a “base year” standard using a set of objects identifiable on all images and deemed to be invariant spectral targets (such as roads, deep water, exposed soil, and sand). Given that the in situ data were collected in 2008, this was chosen as the base year. For each invariant target a mean spectral value in all bands was extracted and matched using a simple regression approach which was then subsequently applied to the 2007 and 2010 images.

Visual inspection of the crown dGPS delineations on the imagery confirmed that the crowns were all well located. In a small number of cases crown locations were aligned to better fit the imagery. The in situ crown condition assessment data were collated and spatially linked to the crowns in the imagery and crown perimeters formed a tree crown mask. Spectral values within each tree crown were then extracted for calculation of the indices.

### 2.4 Index Generation

#### 2.4.1 Textural, spectral, and spatial indices

A total of 125 indices were selected based on (1) the available spectra, (2) literary precedence, and (3) the open source availability of the algorithms used for the gray level covariance matrix textures (GLCM) and gray level difference vector textural indices (GLDV). The GRASS r.texture package was used to generate the textures.36,37

Indices fit into three broad categories of spectral (Table 2), textural (Table 3), and geometric (the area displaced by the crown by counting the number of pixels inside the delineated crown area). The VI used in this study (Table 2) fall into two broad categories (1) narrow-band versions of the normalized difference vegetation index (NDVI) and (2) simple ratios.

Textural indices were calculated from the covariance between neighboring gray levels in an image, resulting in a pixel score of relative positional differences. Rectilinear polygons were formed around each object with a one pixel buffer to perform the GLCM and GLDV to ensure all pixels in the crown were used. Each textural index was calculated for each object in the four cardinal directions. Each resulting textural index was then cropped and all four directions and averaged by band.

Once the individual indices were computed, a mean for each was calculated for the 80 tuart crowns, and this dataset formed the basis of the modeling described below.

Figure 2 provides an example of a true color DMSI image with four tree crown delineations overlaid and an example of two texture measures for an individual crown. The figure shows the marked variation within each crown, as well as the variation through time.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Details of the multi-year acquisition flights made by SpecTerra Services.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>Solar angle</td>
</tr>
<tr>
<td>May 08, 2007</td>
<td>39.18 deg to 40.08 deg</td>
</tr>
<tr>
<td>June 06, 2008</td>
<td>34.01 deg to 34.60 deg</td>
</tr>
<tr>
<td>June 02, 2010</td>
<td>34.74 deg to 35.03 deg</td>
</tr>
</tbody>
</table>
The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.

### 2.5 Modeling Approach

The random forest (RF) algorithm\textsuperscript{26,27} is an ensemble classifier that fits decision trees (models) to subsamples of training data, then combines predictions from the resulting trees. The algorithm bootstrap samples the training data and approximately 63\% of the original observations are used at least once in model development. Data not included in the bootstrapped training sample are called out-of-bag observations. In addition to the randomness of the bootstrapping, the number of variables sampled at each decision tree node is randomly selected from a subset of predictors; by default the number tried (mtry) is the square root of all available predictors.\textsuperscript{27} In this study, mtry is parameterized using the tuning function provided in the \textit{R} implementation of RF by Liaw et al.\textsuperscript{38} and Prasad et al.\textsuperscript{39} as is commonplace for RF applications with many predictor variables. The variable importance score is used to assess the contribution made by each of the spectral, textural, and vegetation indices used in this study.
Variable importance in RF is the normalized difference classification accuracy of each predictor, both when it is included as an observation and when it has been randomly permuted.\textsuperscript{26,27} RF was chosen because it has been shown to have superior predictive power in a number of classification studies\textsuperscript{39–43} and validated against a number of statistical modeling approaches for ecological\textsuperscript{43} and vegetation health modeling applications.\textsuperscript{8}

The 2008 vegetation, texture, and geometric indices were compared to a range of stratifications of the combined TCHI (data not shown). TCHI stratifications of less than 10 point precision (i.e., 10/100) were not considered robust given the human error inherent in the underling crown condition assessments. Two classifications of the \textit{in situ} crown condition indices were used in the final analysis. The first reduced the THCI score to 10\% increments, resulting in a 10 class variable. This level of aggregation is consistent with the estimated error of the measurement protocols discussed earlier. We recognize that as a practical management tool most forest condition assessment indices are applied across a five class scale, so a second analysis was undertaken which reduced the THCI score to 20\% increments to give five classes.

To understand the robustness of the developed models, the relative size of the model development versus model testing datasets was varied from 10\% to 100\% with 10\% indicative of a model which was developed using 10\% of the data and tested on the remaining 90\%, and a 100\% model indicative of a model developed using all observations. Each RF was run 100 times for each split and the out-of-bag error (OOBE), and weighted Cohen’s kappa coefficients ($\kappa$) were calculated (all were significant $p < 0.01$), comparing the models predictions against the observations. $\kappa$ is a statistical measure of inter-rater agreement between categorical data and accounts for a hypothetical probability of chance.\textsuperscript{44,45} Accordingly, $\kappa = +1$ represents

---

**Fig. 2** Map of site 8 featuring (a) the delineated tree crowns, and (b) and (c) two examples of textural crown matrices from which mean crown textural indices were derived. The example indices shown for crown 801 are (b) correlation (COR) for the 780 nm (NIR) band and (c) contrast (CON) for the 450 nm (BLUE) band and each group contains an image from 2007, 2008, and 2010.
complete agreement between observations and predictions and $\kappa = -1$ suggests there is no more agreement other than one would expect by chance. The Landis and Koch\textsuperscript{46} $\kappa$ threshold scale (Table 4) provides a benchmark for interpreting the magnitude of $\kappa$ in the context of the predictive strength.

Once the RF model was developed and assessed, the 100% data model was applied to the 2007 and 2010 imagery and changes in the crown condition index assessed.

3 Results

The $\kappa$ scores were captured for the two models (Fig. 3). A score is given for each of the 10 training and validation fractions (10% to 100% in 10% increments) of the available data ($n = 80$ tuart crowns across the 20 sites). As expected, $\kappa$ increases as the fraction of data used for model development grows larger. Figure 4, a companion to Fig. 3, shows the corresponding OOBE scores for the two models and expectedly, OOBE becomes smaller as the fraction of data used increases. In order to determine an ideal mtry parameterization, the trainRF() method\textsuperscript{38} was used, and this resulted in mtry = 6 for all models except the 10 class 30% and the 20 class 90% and 100% groups which were mtry = 3.

Analysis of variable importance extracted from the 2008 100% training models (Fig. 5) shows that the spectral and textural indices containing the 780 nm spectral band were most often selected in model development. The geometric index of the size of the crown was not a top 5 variable selected by any model. Generally, spectral indices containing the 675 and 780 nm bands were dominant. Although less frequently selected, the 450 and 550 nm spectral bands were also important in both models.

Figures 6 and 7 show the distribution of remote-sensing-derived THCI values of the 80 crowns for 2007, 2008, and 2010 using all the available 2008 data for model development. A shift of the distribution to lower classes (smaller number) is indicative of a decline in crown condition, whereas conversely a shift of the distribution to healthier classes (larger number) indication of a recovery of condition. For example, Fig. 6 shows some recovery between 2007 and 2008, with 10% of the crowns populating the healthiest class in 2008 that were not classified as class 5 in 2007.

Between 2008 and 2010 crowns contracted from both the least and most healthy classes to populate moderate health classes. The net result could be considered a recovery with the majority of crowns healthy, 60% were classified as class 4 (5 class model refer Fig. 6) and class 7 (10 class model refer Fig. 7) which represents a change of almost 20% from 2008 to 2010. Indeed, the least healthy classes 1 (Fig. 6) and 1 and 2 (Fig. 7) show continuous recovery between 2007 to 2008 but not complete recovery over this period. This suggests a recovery from disturbances (prior to 2007) can take a number of years to manifest. Conversely, the increased frequency of moderately healthy classes contributed by both additional healthy and declining crowns is indicative of the resilience of the species to dealing with, and recovering from dieback over many years.

<table>
<thead>
<tr>
<th>Weighted Cohen’s kappa</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>−1.00 to 0.00</td>
<td>Poor</td>
</tr>
<tr>
<td>0.00 to 0.20</td>
<td>Slight</td>
</tr>
<tr>
<td>0.21 to 0.40</td>
<td>Fair</td>
</tr>
<tr>
<td>0.41 to 0.60</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61 to 0.80</td>
<td>Substantial</td>
</tr>
<tr>
<td>0.81 to 1.00</td>
<td>Almost Perfect</td>
</tr>
</tbody>
</table>

Table 4 Weighted Cohen’s kappa coefficient strength of agreement benchmarks after Landis and Koch.\textsuperscript{46}
The remote sensing driven eucalypt crown condition indicator presented here provides insight into the extent and severity of tuart decline in Yalgorup National Park between 2007 and 2010. Crown condition shifts from the less healthy classes to healthier classes (2007 to 2008) and retracts back from the healthiest classes to moderately healthy classes (2008 to 2010). In the last year of observation, 2010, the majority of tuart (60%) are predicted to be of moderate health. These results are indicative of a combination of mild stressors impacting the condition of the tuart, a finding supported by recent evidence of pathogens and soil bacterial communities, both slow-acting agents, being linked to these declines. Eucalypt crown decline often occurs over a number of years rather than instantaneously. However, severe weather, in particular drought stress, pests, and some diseases can cause rapid defoliation events from time to time. Generally, the condition of the dominant species is a surrogate for ecosystem condition health with crown condition becoming widely used as an indicator in land management policy for describing aspects of aesthetics, production, and the biodiversity of ecosystems. Remote sensing driven techniques are increasingly useful in providing data to support the development of these critical indicators.

These results confirm an enhancement of the eucalypt crown condition indicator from Evans et al. and the RF variables importance results (Fig. 6) indicate how this was done. Spectral and textural indices containing the 780 nm spectral band were most often selected, an expected result given the evidence that near infrared reflection is sensitive to changes in foliage condition. All of the VIs most commonly selected by the models utilized the 780 nm band in addition to 675 nm (red) which is a strong indicator of chlorophyll absorption. VI containing 450 nm (blue) and 550 nm (green) were also important, and this suggests that variations in this part of the electromagnetic spectrum, for example green reflectance and blue absorption, are being driven...
by changes in foliage condition. This strong relationship between the blue and green spectral bands is known to explain variance in crown chlorophyll content\textsuperscript{51-56} and as a result of a shift in the location and position of the red edge from the red-nir to blue-green part of the electromagnetic spectrum as chlorophyll content decreases from leaf senescence or absence.\textsuperscript{24,57,58}

An interesting finding of this study is that as the number of increment classes changed from small (5 class) to large (10 class) the tendency of the RF to select texture-based indices increased. This result implies that texture-based measures allows for finer discrimination of more levels of crown condition whereas spectral-based indices allow for more broad scale canopy assessments.

The RF modeling approach used in this study allowed us to investigate two key aspects of the crown condition dataset. First, as the RF modeling procedure is not computationally expensive, iterative modeling and the bootstrapping approach provides insight into the robustness of the developed models. The results indicate, as expected, that as a higher proportion of data is used in model development the overall accuracy of the model improves in both the 5 and 10 class data-sets. Second, a typical jackknifing of the dataset into 70\% for model development and 30\% for independent model prediction in this study produces $\kappa$ of 0.67 for the 10 class and 0.64 for the 5 class models, a substantial agreement on the Landis and Koch\textsuperscript{46} benchmark (Table 4), and similarly promising to reported accuracies of Evans et al.\textsuperscript{25} who applied standard linear regression of mean crown VI to classify condition. Both models achieve a $\kappa$ of 0.84 when using 90\% of the observations and $\kappa = 1$ when all of the available data are used. It is important to note that given RF internally bootstraps, this means that when the models are developed using only 90\% of the observations RF is actually using as little as 45 crowns.

Whilst RF is not a descriptive or inferential tool it is useful for identifying important ecological variables for interpretation.\textsuperscript{43,59} We have used RF variable importance to isolate useful spectral and textural indicators of crown health through classification and prediction of a categorical set of

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Comparison of the random forest out-of-bag error score for the 10 and 5 class models for each of the training and validation fractions used in the model development.}
\end{figure}
RF is non-parametric, requires no distributional assumptions, and can handle the ‘small n large p’ scenario demonstrated in this study, by the splitting the available data into incremental portions.\textsuperscript{27,38} RF offers a robust, efficient, and accurate alternative to parametric and semi-parametric methods for identifying ecological indicators.

As expected, when comparing the 10 to the 5 class datasets the results indicate the OOBE, a standard error indicator of RF’s internal predictive capacity, is higher for the 10 class when compared to the 5 class model confirming that a smaller number of crown condition classes produces a more accurate overall model.\textsuperscript{27,38,39,60-63} The trade-off between classification precision (number of classes) and model accuracy (OOBE and $\kappa$) is critical to its application as an indicator where it is likely a lower precision classification is sufficient for forest health management, given limitations to response options, and therefore greater accuracy would contribute to greater confidence in its application.

The development of remotely derived eucalypt crown condition indicators is imperative for the management of the large forested estates and to quantify changes in structure and health over time. High spatial and spectral remotely sensed data is a rich source of information for forest managers and as demonstrated in these results, can make accurate predictions of crown condition across a large number of plots and crowns. Critical to the success of this type of remote sensing data is the capacity to explicitly model changes in the biophysical properties of tree crowns, such as a reduction in chlorophyll and changes in crown structure such as defoliation.

In terms of the application of this technique to forest and woodland management the 5 class assessment could offer broad insights to overall canopy condition whereas the 10 class variable
could offer potentially increased sensitivity to internal crown structure and condition. This additional sensitivity would be valuable, for example, when looking for evidence of the initial stages of eucalypt crown dieback and discerning the different agents involved in the declines.

As discussed, in situ crown condition data collection is an expensive task and as a result remote sensing and modeling approaches could represent a potential cost saving through both limiting the number of site visits and reducing the sample size of crowns assessed in the field. We recognize that the assessment of 80 crowns across the national park for model development is relatively small and that, as this work continues, the collection of additional

---

**Fig. 6** Histograms of the total crown health index (TCHI) classification using the 5 class model where 1 = least healthy or declining (TCHI of 1% to 20%) and 5 = most healthy (TCHI of 81% to 100%) predictions for (a) 2007 and (c) 2010; and (b) the in situ observations for 2008. All percentages calculated using all available data (n = 80).

---

**Fig. 7** Histograms of the total crown health index (TCHI) classification using the 10 class model where 1 = least healthy or declining (TCHI of 1% to 10%) and 10 = most healthy (TCHI of 91% to 100%) predictions for (a) 2007 and (c) 2010; and (b) the in situ observations for 2008. All percentages calculated using all available data (n = 80).
in situ data is imperative, especially to test the predictions through time. Evans et al.\textsuperscript{25} and Horton et al.,\textsuperscript{30} however, indicated that information on 80 tuart crowns was a suitable number for the assessment of canopy condition from the ground. In addition the capacity of the RF approach to allow investigation of both the number of classes, and the robustness of the relationships provides confidence in these results.

Finally, the collection of additional in situ data and the use of remotely sensed mean crown statistics in the RF models required labor-intensive and expensive manual delineation of crowns both in the field and within the imagery. Future work will use object-based image analysis (OBIA) software that has been demonstrated as an alternative to delineating crowns manually using processes such as feature extraction or segmentation. This process has been demonstrated on non-eucalypt biomes\textsuperscript{64-66} and eucalypt biomes,\textsuperscript{67} but it has not been conducted at the crown scale on native eucalypts exhibiting advanced crown decline and will therefore present some challenges.

5 Conclusion

This study has demonstrated the application of RF models trained using remotely sensed spectral, textural, and geometric indices to predict eucalypt crown condition and its change through time. The results indicate that two models, using 5 and 10 descriptor classes were accurate and robust and could represent a cost-effective way of monitoring and quantifying crown condition of tuart over large areas within an operational environment. The results confirm that spectral and textural information in the 450, 550, 675, and 780 nm spectral bands was critical in the model predictions and texture indices were more informative when increasing the sensitivity of the models to smaller variations in crown condition. Once validated, the models confirmed the overall condition of tree crowns within Yalgorup National Park declines from 2007 to 2010. This study has shown the potential that exists for the use of this indicator of eucalypt crown condition indicator for monitoring changes in the crown condition of tuart and other eucalypt species in native and plantation forests.

Acknowledgments

We are grateful for the funding from the Australian Research Council (LP0668195) Department of Conservation and Environment, Alcoa of Australia, Mandurah City Council, and Bemax Cable Sands. We thank SpecTerra Services Ltd. for their substantial support in the provision of the orthorectified digital multispectral imagery utilized for this study. We also thank Dr. Frank Honey and Professor Nicholas Coops and the anonymous reviewers for their feedback at various stages. The work was supported by iVEC (http://www.ivec.org) through the use of advanced computing resources provided by the Australian Resources Research Centre Facility located at Technology Park, Perth. The work of Bradley Evans is supported by a Murdoch University PhD scholarship and a Centre of Excellence for Climate Change Woodland and Forest Health scholarship.

References


Queries

1. Corresponding address taken from database. Is it okay?

2. This query was generated by an automatic reference checking system. The references 13, 26, 33, 34, 38, 65, 67 could not be located in the databases used by the system. While these references may be correct, we ask that you check them so we can provide as many links to the referenced articles as possible.

3. Please provide information that tells what kind of reference Ref. 10 is.

4. Please provide publisher name and location for Ref. 18.

5. Please provide publisher name for Ref. 28 and 29.

6. Please provide issue number for Refs. 33, 34, 50, 59, 61, 64.

7. Please provide page range for Ref. 38.

8. Please provide publisher name and location for Ref. 40.

9. A check of online databases revealed a possible error in this reference. (53) The year has been changed from '1951' to '1952'. Please confirm this is correct.

10. Please provide volume and issue number for Ref. 56.

11. This article currently does not include biographies or photos for any of the authors. If you wish to include biographies, please provide text and photos with your corrections.